Visual statistical learning and relearning:
A cohort study with children age 5 to 7 years and adults

Dissertation
zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

an der Universität Hamburg
Fakultät für Psychologie und Bewegungswissenschaften
Institut für Psychologie

vorgelegt von
Daniela Schönberger, M.Sc.

Hamburg, 2023

Promotionsprüfungsausschuss:

Vorsitzende: Prof. Dr. Barbara Hänel-Faulhaber

1. Dissertationsgutachterin: Prof. Dr. Brigitte Röder

2. Dissertationsgutachter: Prof. Dr. Ulf Liszkowski

1. Disputationsgutachter: PD Dr. Patrick Bruns

2. Disputationsgutachterin: Prof. Dr. Jutta Mueller

Tag der Disputation: 06.09.2023

Abstract

Tracking sequential information allows children to quickly acquire environmental statistics which is important for, e.g., language learning. Thus, it has been suggested that children should excel in implicit statistical learning and outperform adults therein. Older children and adults seem to rely more on explicit learning mechanisms. Findings on how the retention and the generalization of sequential regularities changes across development have been inconsistent, however. The employed research designs did not allow drawing conclusions, neither about long term retention, nor about transfer of sequential knowledge.

The present dissertation aimed at closing this gap by investigating three child groups (5-year-olds, 6-year-olds, 7-year-olds) and two independent groups of adults, who learned visual sequences which were defined by complex rules as defined by an Artificial Grammar (AG). This age range was selected due to well-known extensive changes in cognitive functions, such as memory and cognitive control. All groups completed three learning sessions on separate days over the course of one week (Year 1). After one year, remaining AG knowledge was tested in three “relearning” sessions with the original item set (Year 2). In a subsequent session, transfer to a new visual stimulus set was tested which implemented the same AG. Seven-year-olds and one adult group served as controls for the other groups after the delay, so their study design differed slightly. They both had an additional transfer session at the end of Year 1, to test rule generalization before the delay. Their learning trajectories were discussed separately as Project 1, while learning outcomes of 5-year-olds, 6-year-olds and the second adult group were compared in Project 2. Project 1 and Project 2 both included assessments of explicit sequence knowledge and memory and German grammar skills in children and adults, which were related to AG learning for all participating groups in Project 3.

We hypothesized to find an increase in learning performance across sessions in all child groups and adults for the first stimulus set in Year 1. We expected to observe preserved AG knowledge as well as transfer effects in all groups. Younger age groups were predicted to feature higher retention over one year and quicker relearning of the AG with the first stimulus set, and to show larger transfer to a second stimulus set in Year 2. Additionally, we assumed that children predominantly rely on implicit knowledge, while adults acquire more explicit knowledge about the underlying AG. We predicted that higher capacities in working memory and declarative memory retrieval, and stronger German grammar skills are associated with better AG task performance.

Children from 6 years onwards successfully learned the visual AG, showing an adult-like increase in learning across three sessions while being continuously outperformed by adults. All age groups used their acquired AG knowledge after a 12-month delay for quicker relearning of the same input that enabled them to reach higher final performance levels, as compared to Year 1. Furthermore, all age groups transferred their AG knowledge to new surface features. However, relearning results did not confirm that younger children outperform older children and adults in the long run (Project 1 & Project 2). When controlling for maturational effects after the delay in 5-year-olds and 6-year-olds, both groups showed quicker re-acquisition of AG rules in the first session compared to age-matched naïve controls. The group who started at 5 years of age showed gains from prior learning when they were relearning at age 6 years, even though they had not demonstrated successful learning of the AG in the first year (Project 2). Levels of explicit AG knowledge did not differ between any of the investigated age groups (Project 1 & Project 2). Exploratory evidence was provided that memory skills (working memory & declarative memory encoding/retrieval), and to a lower degree German grammar skills, were associated with multi-session AG learning outcomes (Project 3).

The present findings corroborate the idea that repeated exposure to sequential regularities results in long-lasting memory traces and the ability to generalize these regularities to a situation with new visual input. However, the current study does not provide support for superior retention of sequential regularities over a longer time period early in development, neither for younger children as compared to older children, nor for children as compared to adults.

Contents

Abbreviations

ACS	Associative Chunk Strength
(rm)ANOVA	(repeated measures) Analysis of Variance
AG(L)	Artificial Grammar (Learning)
BA	Brodmann Area
BF	Bayes Factor
BSCL	Brief Symptom Checklist
CBCL	Child Behavior Checklist
CI	Confidence Interval
CLS	Complementary Learning Systems
D-PA	Deutschtest für die Personalauswahl
ERP	Event-related Potential
KABC II	Kaufman Assessment Battery for Children II
K-TIM	Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene
MEG	Magnetoencephalography
MultiPic	Multilingual Picture Database
PFC	Prefrontal Cortex
SET	Sprachstandserhebungstest
tDCS	Transcranial Direct Current Stimulation
TMS	Transcranial Magnetic Stimulation
WAIS-IV	Wechsler Adult Intelligence Scale IV

Chapter I: General Introduction

Children are excellent learners, who seem to quickly acquire new skills like language without much effort and despite the fact that their general cognitive abilities are still developing. By age four to five years (Werker & Hensch, 2015), they master the grammar of their native language, even though they are able to store and update only a limited amount of information in their working memory (Gathercole, 1998), are poor at remembering specific episodes or facts (Gathercole, 1998; Keresztes et al., 2017), and struggle at switching between different task demands or inhibiting prepotent responses to irrelevant input (Hughes, 2013; Ramscar & Gitcho, 2007). A large body of research has addressed the question, which environmental factors and neurocognitive mechanisms contribute to making childhood a time of effective and adaptive learning (Aslin, 2017; Gualtieri & Finn, 2022; M. H. Johnson & Munakata, 2005; Ramscar & Gitcho, 2007). A potent cognitive mechanism that allows learners to effortlessly extract environmental regularities has been identified from studying language development: Saffran et al. (1996) were the first to show that infants use frequency information from co-occurring syllables to learn words from continuous speech input. This mechanism was termed "statistical learning" and subsequent research has since provided evidence that it operates not only on auditory/language input, but allows the extraction of visual, visuomotor and tactile patterns from the environment (Conway, 2020). Furthermore, statistical learning mechanisms are available to learners of a wide age range from infancy to adulthood (Aslin, 2017; Conway, 2020), and aid the acquisition of more complex – e.g., non-adjacent – regularities (Mueller et al., 2018; van der Kant et al., 2020), which are important for skills like grammar learning (Uddén & Männel, 2018).

The tracking and use of environmental regularities is most often discussed in the context of language learning, e.g., how we learn syntax rules and segment speech input into words. This mechanism has been argued to rely on "statistical learning" (computing transitional probabilities between co-occurring items) and, closely related, "implicit learning" (chunking co-occurring items in memory) operations (see Conway & Christiansen, 2009 and below, on reconciling both literatures): Performance in tasks that require learning sequential regularities was not only shown to be cross-sectionally related to language processing abilities in typically developing populations (Conway et al., 2007; Misyak & Christiansen, 2012; Misyak et al., 2010; Smith et al., 2015), but also seems to predict later language impairments, such as developmental dyslexia ("procedural deficit hypotheses"; Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pierpont, 2005). At the neural level, this link has been substantiated by fronto-parietal "language areas" (e.g., Broca's region, BA 44/45) being recruited for learning complex sequence rules from a rule set, called “Artificial Grammar” (Conway & Pisoni, 2008; Goranskaya et al., 2016; Skosnik et al., 2002). Stimulation studies have implied a causal role of these “language networks” in learning sequence rules. They report that task performance in discriminating rule-following vs. rule-violating sequences was successfully manipulated by targeting core areas of this network like Broca’s region via current (tDCS) or magnetic (TMS) stimulation, respectively, at test (Uddén et al., 2008; Uddén et al., 2017; Vries et al., 2010). Above and beyond language, extracting sequential regularities from the environment was furthermore reported to influence how we segment our continuous everyday experiences into discrete events and predict upcoming events (Levine et al., 2019), and even how we update representations of associated objects in memory (Yu & Zhao, 2018). Thus, it has been established that learning sequential regularities is important for skills that are acquired over a longer developmental time. However, investigations that systematically test how the developmental timing of learning sequential regularities influences learning outcomes in the long run are largely missing.

To set the stage for a cohort study that aims to fill this gap, we embrace an integrative perspective on the literature of implicit learning and statistical learning and review evidence from both approaches. The present work shares the view of other researchers in this field (Christiansen, 2018; Conway, 2020; Conway & Christiansen, 2009; Fiser & Lengyel, 2022; Pavlidou & Bogaerts, 2019; Perruchet, 2019; Perruchet & Pacton, 2006) that a combined review of both literatures speaks best to the question of how environmental patterns are extracted and used across development, and provides valuable insight into which cognitive and neural mechanisms might underlie this learning process. In the following, I will refer to this learning process as sequence learning.

1. Learning and (brain) development

Sequence learning is evident from very early on in infancy, as demonstrated by newborns’ neural markers like event-related potentials (ERPs) showing sensitivity to co-occurring auditory input (Fló et al., 2022; Kudo et al., 2011), and continues to be effective until late adulthood (Conway, 2020); however, developmental differences have been found in (1) the situations in which learners pick up environmental regularities (passive exposure vs. task relevance), (2) their ability to behaviorally discriminate regular input from rule-violating variations of this input (i.e., performance in two-alternative forced choice tasks), and (3) the information represented in memory after exposure to encountered regularities (transition specific vs. group/broad representations) (reviewed in Forest et al., 2023).

1.1. Determinants of age differences in sequence learning

Previous literature reviews (Aslin, 2017; Conway, 2020; Daltrozzo & Conway, 2014; Forest et al., 2023; Krogh et al., 2013) have identified several dimensions for determining if and how measured outcomes in sequence learning change with age. These will be elaborated below, based mainly on Forest et al. (2023), who presented the most recent and comprehensive review.

Using indirect learning markers (term by Forest et al., 2023) like ERPs for rule-following vs. rule-violating sequences (Daltrozzo & Conway, 2014) and neural entrainment to sequentially structured stimuli (e.g., reflected in inter-trial coherence indexing event-related phase locking; Choi et al., 2020), infants' and children's sensitivity towards passively encountered regularities has been well documented. In contrast, adults seem to track sequential regularities more selectively depending on their task relevance: Rohlf et al. (2017) showed that only infants' ERPs evidenced learning of cross-modal regularities during passive exposure, while adults' ERPs differentiated between regular and newly combined stimulus pairings only when they were task-relevant. Mueller et al. (2018) identified the age range between two and four years as the optimal time in development for linguistic input (syllables) following non-adjacent sequence rules, after which tracking regularities from mere exposure is not observed to the same degree as earlier in development. Similarly, in later childhood, indirect behavioral measures of "online" learning paint a picture of an early sensitivity towards sequential regularities: These measures include improved reaction times to high-probability vs. low-probability sequence locations, which were reported to reflect a high sensitivity for visuomotor regularities in the youngest age group of this study (age 4 to 12 years) that decreases in adolescents and adults (Janacsek et al., 2012). Indirect learning markers comprise measurements during both exposure to rule-following input (e.g., online tracking of regularities in EEG markers or reaction time improvements), and at a later phase (e.g., ERPs or looking times in response to legal vs. illegal sequences; Forest et al., 2023). To sum up, indirect learning markers suggest successful sequence learning at all ages, with a heightened sensitivity towards passively encountered or implicitly embedded regularities in infants and young children vs. older children and adults.

A different age-pattern emerges when sequence learning is measured as discrimination performance between regular vs. irregular input at test, after having been exposed to reoccurring regularities: These more direct learning markers (term by Forest et al., 2023) show that discrimination improves from 5 to 12 years (Raviv & Arnon, 2017;

Shufaniya & Arnon, 2018) and, extending this age range to adults, from 6 to 30 years (Schlichting et al., 2017) in visual sequence learning. The distinction between developmental trajectories in indirect (sensitivity in tracking regularities) and direct (displaying sequence knowledge at test) learning markers can accommodate both the reports of a high initial sensitivity for statistical regularities which decreases later in childhood (in indirect learning markers, see paragraph above) and an improving ability with age to tell apart input which follows vs. such which violates the learned sequential and probabilistic relationships (direct learning markers; see also Lukács & Kemény, 2015; Weiermann & Meier, 2012 for more skill-based and probability learning tasks).

With regard to the information represented in memory after sequence learning, it has recently been suggested that there is a developmental shift in how general vs. specific information about encountered regularities is represented in memory (reviewed in Forest et al., 2023). Recent evidence shows that younger children (5-7 years old) appear to remember only item-specific transitions from learning three-item sequences (triplets), while older children (8-9 years old) and adults represent specific (item-level transitions) and broad information about triplets (group membership) in memory (Forest et al., 2021; explained in detail below). This study pitted different levels of sequence information in one input stream against each other to directly test how representations vary by age. Still, children (and for that matter, even infants, see M. C. Frank et al., 2009; Marcus et al., 1999) have been shown to extract higher-order information like category-level rules when exposed to item-level regularities (Jung et al., 2020; Nowak & Baggio, 2017). I will try to reconcile these findings in the context of memory development later in this section and when taking a closer look at generalization (section *Transfer in sequence learning*).

In summary, age differences in sequence learning depend on the learning marker (direct vs. indirect) as well as the exposure situation at hand (passive vs. task-relevant). Apart from these qualitative changes in sequence *learning* across age, the type of information that is represented in *memory* from a sequence learning experience (e.g., specific item-level transitions vs. broad/group membership) might change as a function of age as well.

1.2. Sensitive periods and neurocognitive mechanisms underlying sequence learning

Ongoing brain development plays an important role in how learning measures differ depending on age; in particular, it shapes which learning mechanisms are available and predominantly used, and constrains what is represented in memory from learning experiences at a certain point in development. It has been argued that there are restricted time windows during development, “sensitive periods”, during which the environment exerts a stronger influence on brain and behavior than at other times (Knudsen, 2004). This has been demonstrated in the context of higher cognitive functions; for instance, early language exposure shapes language acquisition and its underlying neural circuitry in the long run (Werker & Hensch, 2015). In their review, Werker and Hensch (2015) describe “switch” mechanisms, mainly the inhibitory/excitatory balance of parvalbumin cells and molecular brakes, that render brain structures and functions easily and extensively malleable by environmental input during limited time periods in development. Changes in the neural infrastructure during that time then shape later learning experiences that use and build on what has already been established. This provides a scaffold for probable future input that is similar to the early encountered experiences, but at the same time constrains the capacity to extensively adapt to completely new learning environments or input (Knudsen, 2004). While at the neural level, this mechanism has been mainly studied for auditory and visual adaptations in non-human species (Keuroghlian & Knudsen, 2007; Knudsen, 2004), there also seem to be long-lasting behavioral consequences from early learning experiences in human speech perception: Speech sound (phonetic) discrimination from a native language that had not been used after infancy was successfully retrained later in adulthood (Werker & Hensch, 2015). Additionally, early reports in second language acquisition (J. S. Johnson & Newport, 1989), suggested that native-like grammar skills in adulthood can only be achieved through exposure to grammar input before age 7. This idea of time periods in development with an increased sensitivity for certain input has been extended (Gualtieri & Finn, 2022) to other domains, such as sequence learning: Janacsek et al. (2012) showed that children between 5 and 12 years improve to a greater degree than older age groups in their reaction times to predictable vs. random sequence positions in a visuomotor sequence task. Based on this, the authors argue that skills that rely on exploiting (simple) sequential regularities might best be acquired before the age of 12 years, which maps onto the time course of second language learning (Gualtieri & Finn, 2022; Qureshi, 2016).

In addition to circumscribed periods in life that have been shown to shape neural functioning and learning behavior in the long run, neurocognitive accounts propose changes in the relative contributions of learning systems with age: Several authors have organized the age-related findings mentioned before by stressing that sequence learning mechanisms shift from more implicit to more explicit processes across development (Conway, 2020; Daltrozzo & Conway, 2014; Janacsek & Nemeth, 2012; Nemeth et al., 2013). In this view, an implicit learning system, relying on modality-specific neural systems governs learning early in development. It helps to extract simple patterns in a “model-free” and data-driven manner that requires little attentional resources. Over the course of later childhood and adolescence, an explicit learning system takes over, which recruits domain-general neural systems (presumably based on prefrontal cortex [PFC]) to extract complex patterns by leveraging attentional resources in a goal-directed manner. This “model-based” learning system then continues to be the dominant mechanism recruited in learning situations across adulthood. Support for this model (e.g., Jost et al., 2011) comes mainly from studies that look into early vs. late ERP components (reviewed in Daltrozzo & Conway, 2014). It is important to note that early-developing implicit mechanisms might interact with emerging explicit components, e.g., attentional resources, in a manner that puts children at an advantage in certain learning environments. For instance, an impoverished ability to allocate attentional resources allowed 7- to 9-year-old children to learn irrelevant information in a divided attention task that looked into the learning sensitivity (earlier identification of old vs. new fragmented drawings) for previously seen, task-irrelevant pictures (Tandoc et al., 2022): In contrast to adults, whose learning advantage from the previous picture presentation decreased in a divided attention setting, children’s sensitivity towards task-irrelevant information remained unaffected. Remarkably, they did not experience a trade-off for task performance on the simultaneously presented goal-related information, when task demands were matched in difficulty for both age groups (see, S. M. Frank et al., 2021 for a similar phenomenon in visual perceptual learning in children vs. adults).

Studies have tried to elicit this childhood advantage in adults by manipulating the balance of implicit vs. explicit learning mechanisms experimentally: Further promoting explicit learning strategies by task instructions in adults (“effortful” learning) was reported to be detrimental for extracting grammatical categories as compared to passive listening (Finn et al., 2014). On the other hand, making adults rely on more child-like learning mechanisms by decreasing cognitive control (i.e., having them engage in a dual working-memory task) prior to passive exposure to hidden auditory regularities, fostered implicit word learning (Smalle et al., 2022). The recruitment of neural mechanisms supporting implicit vs. explicit learning has been directly targeted as well: Disrupting brain areas that usually contribute to explicit learning in adults by transcranial direct current stimulation of the dorsolateral PFC (BA 9) during the acquisition of statistical regularities caused adults to engage more of the implicit learning mechanisms used by children (Friederici et al., 2013). When the same brain area was deactivated in adults during language exposure, word-form learning (Smalle, Panouilleres, et al., 2017) and word segmentation performance were found to be enhanced (Smalle et al., 2022). Applying this stimulation set-up during learning before a 24-hour delay, later (re)learning of visuomotor regularities was reported to be improved as assessed in a second learning session after the 24-hour consolidation period (Ambrus et al., 2020). This suggests that benefits from implicit learning mechanisms might extend to longer timescales and can result in improved performance after a delay period.

To sum up, the literature on sensitive periods in learning proposes that early learning shapes later learning and that children might exhibit a heightened sensitivity for sequential regularities before age 12, with the time window for most efficient acquisition possibly closing as early as age 7. With regard to the extent that implicit vs. explicit mechanisms are recruited during learning, the respective balance seems to shift towards more explicit learning over development. Relying more on implicit processes, as (younger) children do naturally, can be advantageous in situations where a broad attentional focus benefits learning outcomes (i.e., for learning from passive exposure in language, and for picking up task-irrelevant information) and possibly for later use of implicitly acquired sequence knowledge after a delay.

1.3. Memory development and sequence learning

Not only do the brain mechanisms *during learning* change as children get older; the memory representations formed during learning, i.e., the *result of learning*, also seem to differ across development as briefly stated earlier. Forest et al. (2023) proposed that representations from statistical learning are generally implicit memory traces that share properties of other implicitly acquired memories (slow acquisition, context specificity and long-term durability). As briefly mentioned earlier, Forest et al. (2021) assessed memory representations in two groups of children by testing their ability to identify several types of sequence rule violations after exposure to a stream of triplets (i.e., three-item sequences). Younger children, aged 5 to 7, seemed to represent only specific information about item-to-item transitions. In contrast, older children (age 8-9 years) and adults exhibited broader representations that included group-level information. Compared to older children and adults, young children were more likely to correctly identify a sequence of reordered triplet items as unfamiliar. Older children and adults, however, embraced these sequences, which displayed a novel order of the correctly grouped items making up a triplet, more often as familiar (Forest et al., 2021). These age-related differences in representations could be related to the developmental time courses at which neural memory systems become fully functional (described in the following based on Forest et al., 2023), determining which memory processes prevail at different ages.

The maturing hippocampus and its connections to (frontal) cortical areas have been described as the main drivers of age-related changes in how information from sequence learning is represented in memory (Keresztes et al., 2018; Schlichting et al., 2017). In infants, hippocampal learning was described to rely mainly on the early-maturing monosynaptic pathway (Gómez, 2017; Lavenex & Banta Lavenex, 2013; Schapiro et al., 2017). In addition, cortical learning systems seem to be available early on and support association learning across several encounters (McClelland et al., 1995; O'Reilly et al., 2014). Both learning systems have been proposed to require repeated exposure and learn slowly, suggesting that representations early in life lack specificity and stress commonalities rather than differences between learning experiences. From around 24 months of age onwards, the trisynaptic pathway of the hippocampus emerges (Keresztes et al., 2018; Lavenex & Banta Lavenex, 2013; Ribordy et al., 2013) and seems to enable increasingly specific representations of encountered information, which have been suggested to guide behavior in early childhood up to 6 to 7 years of age (Forest et al., 2023). At the same time, the protracted development of the hippocampus' subfields along its anterior-posterior axis was proposed to shift memory processes from mainly generalization (pattern completion) to increasing specificity (pattern separation) between the ages of 4 and 6 years (Keresztes et al., 2018; Langnes et al., 2020; discussed in detail in section *Predictions from memory theories for age-differences in transfer*). With regard to hippocampus-cortex interactions, connections of the hippocampus to the inferior frontal gyrus (likely subserving specific encoding of experiences) were reported to develop prior to those between the hippocampus and medial prefrontal areas (possibly supporting broader properties of an experience or schematic knowledge; Calabro et al., 2020; discussed in Forest et al., 2023). Starting around the age of 7 years, increasingly broad representations of sequential regularities, entailing memory specifics alongside schematic information, are supported by protracted development of the anterior hippocampal subfields (DeMaster et al., 2016; Langnes et al., 2020; Lee et al., 2020) and its later emerging connections to medial PFC (Calabro et al., 2020; put forward by Forest et al., 2023).

This proposed non-linear trend in sequence memory representations (non-specific but rather inflexible in infants, increasingly specific up to roughly age 7, broader including specific and higher-order information from age 8 onwards), is further elaborated in Forest et al. (2023). They caution that it needs to be further put to test empirically.

The later use of sequence knowledge is not only determined by the kind of representations formed from learning experiences, but also depends on how representations can be accessed, retrieved and used; this puts a focus on how memory processes interact with prior knowledge and cognitive control (Brod et al., 2013), both of which are arguably reduced in (younger) children as compared to adults. It has been suggested that prior knowledge affects how memory processes in the medio-temporal lobe (binding episode features into a coherent memory representation) interact with those in PFC (strategic control at encoding and retrieval), thereby accounting for age-dependent changes in relational memory (Brod et al., 2013). In their study on associative learning and memory integration, Shing et al. (2019) suggest that information across related experiences is stored separately in middle childhood (children aged 9-10 years). These experiences are integrated late in the memory process by making inferences at retrieval, while adults may perform this integration as early as encoding.

In summary, memory *representations* after sequence learning have been proposed to become increasingly specific until middle childhood, before turning into broad representations which incorporate both specific and general aspects of the encountered regularities – a dynamic driven by hippocampal development and its connections to frontal areas. In addition, memory *processes*, as in encoding and retrieval processes, seem to shift from emphasizing generalization in early childhood to stronger memory specificity until middle childhood (predictions from this approach for generalization, see section *Transfer in sequence learning*). Continued development of cognitive control and prior knowledge further influence how information from several learning experiences is integrated into memory, possibly from late (retrieval) in middle childhood to early (encoding) memory stages in adulthood.

2. From single session acquisition to multi-session (re)learning of regularities

Building on the previously reviewed literature on sensitive periods, memory, and age-dependent representations formed from exposure to sequential regularities, a longitudinal perspective on sequence learning seems vital (Arciuli & Torkildsen, 2012). While there is extensive work looking at how different age groups track and acquire sequential regularities in a single experimental session, it is less clear what their learning trajectories look like in the long run: How are previous learning experiences used and built upon later, after several weeks or months? And how does this retention and relearning of encountered regularities in behavior differ as a function of age? Approaches taking a long-term perspective are needed to explain real-life learning of skills (Arciuli & Torkildsen, 2012), e.g. acquiring a grammar in language acquisition, by modeling learning across multiple learning situations over time.

2.1. Learning savings in behavior and neural candidate mechanisms

Early studies on relearning (Livosky & Sugar, 1992; Parkin & Streete, 1988) continued in the tradition of Ebbinghaus, who had observed faster and more efficient learning of familiar as opposed to new lists of nonsense syllables (Ebbinghaus, 1880). These studies used “savings” in the relearning of visual material (less time or effort needed at relearning due to prior learning of the same material) after a delay as a proxy for what they call implicit memory and compared these savings between different age groups. They report conflicting results on how relearning differs depending on age after intervals up to two weeks: Livosky and Sugar (1992) observed an age gradient (3-year-olds > 5-year-olds > young adults) with the youngest age group showing the greatest advantage in learning familiar relative to new picture pairs after two weeks’ time. Conversely, Parkin and Streete (1988) found comparable relearning benefits for 3- to 7-year-old children (3 groups: 3-year-olds, 5-year-olds and 7-year-olds) and adults in a recognition task with fragmented pictures after both a one-hour and a two-week delay, when controlling for baseline performance differences between children and adults.

As a candidate mechanism underlying relearning advantages, non-human animal studies (Xu et al., 2009; Yang et al., 2009) have found persistent structural changes in the cortex (i.e., proliferation and reactivation of dendritic spines) when rodents acquire new sensory or motor skills and relearn them after a long-term delay of several months. Faster and more efficient relearning as compared to new learning could thus arise from the fact that the previously acquired task-specific neural infrastructure can be (re)used more efficiently after a delay period (discussed in Hofer & Bonhoeffer, 2010). Dendritic spines formed from early learning experiences in development were furthermore observed to live longer than spines formed from learning later in life, as shown by in-vitro studies on hippocampal tissue from rodents (Yasumatsu et al., 2008). Relatedly, on a network level, principles of structural plasticity in response to auditory input have been described to differ between the developing and the adult neural system (Keuroghlian & Knudsen, 2007): In juvenile animals during a sensitive phase, molecular mechanisms like synaptic strengthening/elimination seem to enable the developing system to adapt to learning environments in a swift and extensive manner by mere exposure. In adult animals, learning plasticity reflected in functional adaptations can still be observed, but only if the encountered input was attended to or task-relevant. Taking this idea to long-term effects in human behavior, Werker and Hensch (2015) related neural adaptations from early learning experiences in humans to lasting learning advantages for speech sounds later in life. For instance, they review evidence in adopted children, who later as adults were able to retrain discriminating phonemes of their native language that naïve adults are unable to discriminate.

In sum, relearning on a behavioral level seems to be characterized by faster and more efficient learning as compared to first learning in both children and adults, with a possible advantage for young children. These behavioral advantages at relearning could be subserved by structural adaptations in the brain that seem to happen under less restricted conditions (mere exposure), more extensively, and probably with longer-lasting effects early in life.

2.2. Multi-session studies and retention of sequence knowledge

Despite some earlier evidence of preserved knowledge across at least one year in adult visuomotor (Romano et al., 2010), artificial grammar (Allen & Reber, 1980), and artificial language learning (M. C. Frank et al., 2012), only a limited number of studies have tested human retention after a long-term delay in different age groups. These studies looked into the question of how children and adults retain phonological (Ferman & Karni, 2010, 2014; Smalle, Page, et al., 2017) and visuomotor (Kóbor et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021; Tóth-Fáber, Tárnok, et al., 2021) sequential regularities over several months up to one year.

Two multi-session learning studies used auditory syllable sequences to investigate phonological word-form learning (Smalle, Page, et al., 2017) and the learning of an artificial morphological rule (Ferman & Karni, 2010, 2014). Mapping learning trajectories over 4 (Smalle, Page, et al., 2017) to 15 (Ferman & Karni, 2010) sessions, they showed that adults and 8- to 12-year old children acquire a single sequence rule with the same learning rate, despite better overall performance levels in adults. Both studies suggest that learning relies more on explicit knowledge of sequence rules in adults than in children: Adults acquired more explicit knowledge of sequence rules than children and their knowledge level was associated with improved performance and retention (Ferman & Karni, 2010; Smalle, Page, et al., 2017). Children (8-9 years old) showed no such association in one of the studies (Smalle, Page, et al., 2017) and in the second study (Ferman & Karni, 2010), only older children of age 12 (not of age 8) acquired any explicit rule knowledge that subsequently improved their performance. Looking at retention, these studies yielded inconsistent results for learned sequences over delays of several days to months: Smalle, Page, et al. (2017) observed memory advantages in 8- to 9-year-old children for retaining an implicitly acquired syllable sequence up to 12 months after their last learning session. For the longest retention period of 12 months tested in this study, this child advantage reached significance only for matched subgroups with comparable performance levels before the delay. An explicitly cued sequence was retained equally well by both age groups across 4 hours, 1 week and 12 months (Smalle, Page, et al., 2017). Ferman and Karni (2010) investigated learning of a sequential language rule from several training sessions, providing performance feedback, but no instructions about the sequence rule. They observed preserved performance levels and reaction time improvements after a 2-month retention interval for both adults and 12-year-olds, but not for 8-year-olds.

In the visuomotor domain, children aged 9 to 15 years (Tóth-Fáber, Janacsek, & Németh, 2021) and young adults (Kóbor et al., 2017) were shown to retain both frequency-based (statistical) knowledge and order-based (sequence) knowledge over an interval of 12 months from implicit acquisition. Within the group of children, age was not correlated with memory for either statistical or sequence knowledge (Tóth-Fáber, Janacsek, & Németh, 2021). This result suggests that children between 9 and 15 years of age retained the learned visual sequences equally well over a one-year delay. However, the combined interpretation of these two studies (Kóbor et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021) is limited, since they used different experimental designs for children and adults.

To sum up, multi-session studies on auditory sequence learning in children and adults show that children and adults improve across sessions to a comparable degree, despite adults outperforming children overall. Evidence on how development influences the retention of sequence rules is less clear. Some findings suggest that children (age 8-15 years) and adults retain auditory and visuomotor regularities equally well for up to one year when some explicit task component is present (i.e., cueing of sequences at exposure or performance/visuomotor feedback). For long-term retention of implicit sequence knowledge, one study has suggested better retention in children vs. adults. All of this evidence on how learning rates and retention differ across age is limited in the choice of task material and sequence rules: Capabilities for linguistic vs. nonlinguistic stimuli (Raviv & Arnon, 2017; Shufaniya & Arnon, 2018; van der Kant et al., 2020) and adjacent vs. non-adjacent regularities (Uddén & Männel, 2018; Wilson et al., 2020) were reported to follow different developmental trajectories. Hence, memory effects for stimuli outside the auditory domain, and for sequences that implement not only single rues have to be compared across development. More importantly, however, multi-session studies so far have neither implemented more than one follow-up session to map relearning trajectories, nor considered children younger than 8 years. As detailed above (see section *Sequence learning and brain development*), the language and memory literature identified the age range of 4 to 7 as a time period during which learning mechanisms change profoundly. This warrants an investigation of how previously acquired information is used for later learning for children in this age range. Additional motivation to do so comes from the generalization literature, which will be summarized below (see section *Transfer in sequence learning).*

2.3. Sleep and forgetting in long-term approaches to sequence learning

Long-term studies by default include offline periods with sleep in between learning sessions. Lerner and Gluck (2019) have reviewed studies that used a wide range of tasks, all requiring the extraction of “hidden regularities”, among them, statistical and implicit learning tasks. Based on this literature, the authors propose that extracting underlying regularities benefits from non-REM sleep, especially if these regularities are temporal and complex. For visuomotor regularities in motor skill learning, studies with children that systematically manipulated sleep (delay including a nap/night of sleep) vs. no sleep (delay without a nap/night of sleep) between learning sessions have been inconclusive, with some studies even reporting deteriorated performance after intervals with sleep vs. no sleep for children (reviewed in Wilhelm et al., 2012). Janacsek and Nemeth (2012) argue that sleep-dependent improvements in sequence learning require awareness, in that only sequence knowledge that participants of all ages become aware of before sleeping, is consolidated in an offline period with sleep. Notably, studies which systematically investigate the role of sleep on sequence learning have looked at rather short time scales spanning several hours to 1 or 2 weeks, so it is unclear if these effects might simply add up across several offline periods of sleep or what their exact time course across months or years would look like.

Sleep-dependent consolidation of encountered environmental patterns has been suggested to rely mainly on time-compressed replay processes in the hippocampus during slow-wave sleep (Lerner & Gluck, 2019; Wilhelm et al., 2012). These processes were proposed to enable the long-term storage of rule knowledge in cortical networks and consequently the generalization of the acquired rules. As discussed above (section *Memory development and sequence learning*), there is evidence of protracted hippocampal maturation in its different subfields beyond middle childhood (Gogtay et al., 2006; Keresztes et al., 2017; Keresztes et al., 2018), as well as in the hippocampal coupling with (ventromedial/ventrolateral frontal) cortical areas, all of which are implicated in memory formation and consolidation. The interplay of hippocampal reactivation and cortical redistribution of representations underlying consolidation might consequently change across development as well, resulting in age-dependent differences in retention and relearning. Recent structural evidence supports this notion and implies that memory consolidation across offline delays is also related to the brain volume of these areas in both children and adults (Schommartz et al., 2023): 6-year-olds and adults with a thinner medial orbitofrontal cortex were better at retaining object-location associations over 2 weeks, and in the child group, greater (right) hippocampal volume was related to higher retention rates. The authors took this finding to mean that structural brain patterns reflecting greater developmental maturity (more frontal thinning, bigger hippocampus) might underlie better memory integration with prior knowledge.

With regard to accessing representations from prior learning experiences, sleep seems to differentially impact how much explicit knowledge children vs. adults gain about implicitly learned sequential regularities (Wilhelm et al., 2012): Wilhelm et al. (2013) found that children aged 8-11 years benefitted more than adults from a preceding sleep phase when asked to explicitly recall sequential transitions from an implicit visuomotor task.

Faster “decay” or forgetting rates and more forgetting in offline periods between (sequence) learning experiences have been postulated for infants and younger children compared to older children and adults, which could contribute to the predicted age differences in memory representations that were mentioned before (discussed for statistical learning in Forest et al., 2023): While infants might display rather unspecific, general representations from fast and extensive forgetting between repeated exposures to sequential regularities, these representations should shift to being more specific in early childhood (until age 7) due to a slower forgetting rate and maturing memory mechanisms involving the hippocampus and cortex (see section *Memory development and sequence learning*). The information that is represented from several learning experiences at different ages will in turn influence later learning, e.g., in terms of which features of future input children attend to and to what extent they generalize prior knowledge to new input (discussed in the next section *Transfer in sequence learning*).

In conclusion, sleep and forgetting play a role in how sequence knowledge is consolidated over time in children and adults. The (explicit) extraction of temporal rules seems to benefit most consistently from sleep, possibly more so in children than in adults. Brain development factors into this sleep-dependent consolidation, with regions like the hippocampus and frontal areas undergoing age-related changes that likely lead to changes in retention and in the type of representations formed from repeated sequence learning experiences. These learning outcomes are further influenced by higher forgetting rates early in development impacting memory retrieval, which can be predicted to have consequences for generalizing learned regularities to new situations as well.

3. Transfer in learning

Apart from addressing what is learned depending on age from a long-term perspective, it is critical to test whether learning generalizes to new input or learning situations across development. Using knowledge from prior learning experiences in a new context requires that this information is stored in memory on a more abstract level than including only item-features and can be accessed and successfully applied in an adaptive manner. Testing generalization can thus speak to the questions of (1) how abstract the acquired rule knowledge is represented and (2) how flexibly the learning mechanism at play can operate, both of which are central characteristics of learning that are likely to change with age.

3.1. Predictions from memory theories for age-differences in transfer

From a memory perspective, early cortical development (slow learning, based on repetition) as opposed to protracted hippocampal development (fast learning, “one shot” encoding of single episodes) (Kumaran et al., 2016; McClelland et al., 1995), and, more recently, within-hippocampus development along its anterior-posterior subfields (Keresztes et al., 2018), have been suggested to support a shift from generalization to specificity in childhood (Complementary Learning Systems Theory, CLS): In the age range of 4-6 years, processes of pattern completion ("incomplete or degraded representations are filled in based on previously stored representations", Yassa & Stark, 2011, p. 515) that support generalization have been described to decrease relative to processes of pattern separation ("similar representations are stored in a distinct, non-overlapping (orthogonalized) fashion", Yassa & Stark, 2011, p. 515), which start to manifest in a higher memory specificity around age 6, approaching an adult level. In behavior, this age shift has been observed in tasks that tap into relational memory (pattern completion: testing how learned object associations are elicited across similar settings) and mnemonic similarity (pattern separation: testing how well previously seen objects can be discriminated from perceptual lures), implying that children from 4 to 6 years over-generalize less from previous learning experiences and remember more specifically what they had been exposed to across different episodes (Ngo et al., 2018). This finding persisted after controlling for worse general memory capacities in younger children. Additional evidence for greater generalization early in life comes from the language literature, where overregularization of grammatical patterns such as past tense of verb endings (e.g., in English "-ed") in early stages of language learning is a well-documented finding (Marcus et al., 1992). Testing this overgeneralization bias in a controlled study setting, evidence suggests that children – even quite late, at age 6 years – form hypotheses about whether or not to generalize grammatical patterns based on higher level probabilities in the input (Wonnacott, 2011), and that memory limitations in children (vs. adults) are not sufficient to account for this child bias of overgeneralization in a computational model (Perfors, 2012).

Faster and greater forgetting that happens between learning experiences early in development could additionally favor higher transfer abilities in younger children as compared to older children and adults – a prediction derived from the "spacing effect" in category and concept learning (Vlach, 2014) that has been recently applied to statistical learning as well (Forest et al., 2023): The "forgetting-as-abstraction" account (Vlach, 2014) proposes that relevant features of learning items (i.e., shared properties of items belonging to words or categories) that are present at several consecutive learning events are reactivated in memory on these occasions. Their reactivation increases how strongly they can be retrieved and consequently slows down their forgetting rate (forgetting being defined as a "diminishing ability to retrieve information across time", Vlach, 2014, p. 164). In this way, forgetting can strengthen the shared features of learning input over time, which can be transferred to underlying regularities of sequences that can be seen as shared (latent) features. Forgetting rates early in development are very pronounced, as demonstrated for sequential input by (Bauer et al., 2000). All of this would suggest that younger children could show stronger generalization than older age groups across several sequence exposures (discussed as “fuzzy” representations in Forest et al., 2023).

In general, sleep between several learning experiences seems to benefit the extraction of abstract regularities from environmental patterns. As an underlying mechanism, time-compressed replay in the hippocampus during slow-wave sleep (Lerner & Gluck, 2019; Wilhelm et al., 2012) was suggested to enable the long-term storage of rule knowledge in cortical networks, and consequently the generalization of the acquired rules (described in more detail above, see section *Sleep and forgetting in long-term approaches to sequence learning*). This has been further elaborated by a study that looked into transfer-related replay mechanisms in the human hippocampus with magnetoencephalography (MEG) during rest, one day after learning an explicitly instructed sequence rule for remapping item positions (Y. Liu et al., 2019). It reports preliminary evidence that this learned sequence rule was applied to new items, as reflected in observed replay patterns that followed the position rule instead of the actually experienced input sequence. So, new input was reordered in replay according to a previously learned sequence rule, even if this input had never been seen in a rule-conforming order. The authors interpret this sequential replay at rest as one mechanism for generalizing rule knowledge from prior learning to new input.

3.2. Inconsistent findings for transfer in sequence learning across development

Testing generalization in sequence learning, infants as young as 5-7 months have been demonstrated to apply short abstract rules from training sequences to new test stimuli in habituation paradigms (M. C. Frank et al., 2009; Marcus et al., 1999) – with the younger infants only showing learning effects if the rule was presented simultaneously in two modalities (M. C. Frank et al., 2009).

Nevertheless, testing generalization in statistical and implicit learning using behavioral discrimination tasks has provided mixed results for children aged 3 to 12 years in comparison to adults so far. These studies measured generalization as applying learned sequence rules at test to new items of the same perceptual category present during exposure (e.g., new syllables in Artificial Grammar Learning [AGL] or unseen pictures of the same animals in triplet learning): In single-session transfer studies, adults and children aged 3-6 years (Nowak & Baggio, 2017) or 6-9 years (Jung et al., 2020) respectively, have been shown to generalize learned regularities to new items in the auditory domain (Nowak & Baggio,

2017) and to new instances of underlying categories in the visual domain (Jung et al., 2020). Despite successful transfer for children and adults in both studies, they reported that age groups differ in the exact conditions under which transfer effects manifest: Jung et al. found that children failed to express their more abstract rule (category) knowledge under higher retrieval demands (triplet completion task as opposed to transfer in forced-choice tasks). Nowak and Baggio (2017) tested transfer for different rule systems governing the order of sequences (AG) and demonstrated that children perform better than adults in rule systems reflecting distributional patterns from natural languages, but worse than adults in rule systems with non-typological distributions.

Conversely, 10-11-year-old-children altogether failed to generalize a sequential syntax rule to new items in a study using a complex artificial language, while adults succeeded in this transfer task by leveraging the explicit knowledge that they had acquired about this rule (Hickey et al., 2019). A multisession study (Ferman & Karni, 2010) reported that adults and 12-year-olds, but not 8-year-olds, generalized a highly practiced language rule to new items at the end of each of 10 training sessions. Furthermore, only these older age groups (adults and 12-year-olds) retained their last performance level for new items in a retention session after a delay of two months (Ferman & Karni, 2010). Even when provided with five additional training sessions, the younger age group showed no transfer to new items on either timescale (neither at the end of each training session, nor in the follow-up session after two months). Eight-year-old children were only able to transfer their acquired rule knowledge to new items after they were explicitly informed about the nature of the rule, as shown in a follow-up study with a different sample tested in the same task design (Ferman & Karni, 2014). This lack of transfer in (younger) children contradicts predictions derived from the previously introduced memory literature on CLS (Keresztes et al., 2018; Ngo et al., 2018) and the “forgetting-as-abstraction” account (Vlach, 2014). However, it fits well with the finding that younger children (5-7-year-olds) seem to represent transitions on an item-level after sequence learning in memory (specific representations). On the contrary, this study reports that behavior reflecting higher-order (“broader”) representations emerges only around 8-9 years of age (Forest et al., 2021).

To conclude, generalization in (sequence) learning has been hypothesized to be particularly strong early in development. This has been put forward by theories that focus on changes in memory processes and argue that processes shift towards more specific encoding and retrieval around the age of 6 years. In line with this, infants have been reported to transfer simple sequence rules to new stimulus material, as reflected in their looking/listening times. But behavioral evidence in sequence learning which compared transfer effects between children (age 3-12 years) and adults seems inconclusive so far. In the only multi-session study available, children younger than 12 years failed to show transfer effects, except when being explicitly told what constitutes the sequence rule. This finding could be explained by focusing on the represented features of sequential input that might be available to learners of different ages (reviewed in Forest et al., 2023), with younger children displaying only item-specific representations of sequence transitions when presented with foils violating different levels of regularities at test (specific to broad). Thus, memory approaches focusing on encoding and retrieval processes and the "forgetting-as-abstraction" account predict stronger transfer early in development. In contrast, stronger transfer in later childhood is predicted by approaches focusing on features of memory representations. In addition, observed age-differences in sequence learning transfer might be influenced by the access to and use of explicit sequence knowledge, task specifics and the chosen measure of transfer. It is important to note that not only are these findings on transfer effects in sequence learning across development inconclusive, long-term (spanning several sessions or delays > 2 months) transfer studies that include children younger than 8 years are completely missing.

4. The present cohort study

The present dissertation addressed the question of how the developmental timing of repeated sequence learning influences learning outcomes in the long run. To this end, we investigated how three child groups (5-year-olds, 6-year-olds, 7-year-olds) and adults learn visual sequences involving complex rules over several sessions (see Table 1). It was further tested whether children and adults generalize the underlying rules to new surface features in an integrated longitudinal design that spans a one-year delay between two sets of 7 to 8 sessions in total (see Fig. 1 & 8). This extends the existing literature in several ways:

(1) Mapping sequence learning across multiple sessions both before and after a one-year-delay, enabled us to describe protracted learning of children and adults in the long run. Thus, we investigated the relearning of previously acquired regularities in another set of several sessions after one year, not only retention in a single follow-up session, as done before (Ferman & Karni, 2010; Kóbor et al., 2017; Smalle, Page, et al., 2017; Tóth-Fáber et al., 2021; see section *From single session acquisition to multi-session (re)learning of regularities*).

Table 1

A *Study Design for Child Groups with Sessions in Year 1 & Year 2*

Age	5.0 Years	6.0 Years	7.0 Years	8.0 Years
5-Year-Olds	123	456 Transfer 2		
6-Year-Olds		123	456 Transfer 2	
7-Year-Olds			123 Transfer 1	456 Transfer 2

B *Study Design for Adult Groups with Sessions in Year 1 & Year 2*

Timepoint	1	1 + 12 months
Adults 1	123 Transfer 1	456 Transfer 2
Adults 2	123	456 Transfer 2

Note. 1,2,3,4,5,6 = Sessions with stimulus set 1; Transfer 1&2 = sessions with stimulus Set 2, gray = sessions at home.

(2) We included younger age groups than previous multi-session studies (the youngest was investigated in Ferman & Karni, 2010, with age 8 years across a 2-month delay), namely 5-, 6- and 7-year-old children. This age range is of particular interest with regard to the proposed shift in representations from sequence learning and from memory generalization to specificity (see sections *Memory development and sequence learning* and *Transfer in learning*). Our study design further allowed us to control for maturational changes in 5- and 6-year-olds after the delay by contrasting their performance with a naïve control group of the same age, but without prior learning experience (see Table 1). As elaborated in the previous sections, learning mechanisms were suggested to change specifically in language, memory and generalization outcomes between 4 and 7 years of age.

(3) We tested the transfer of learning effects to a new visual stimulus category in a separate session instead of testing transfer to new items from the trained category within the same session as done previously (Ferman & Karni, 2010; Jung et al., 2020; Nowak & Baggio, 2017). This speaks to the extent to which regularities acquired in previous training sessions can be generalized to an unfamiliar stimulus material in a new learning situation, possibly leveraging advantages from sleep and forgetting in offline periods between sessions.

(4) Instead of a single rule governing syllable sequences in the auditory domain as implemented in previous multi-session studies (Ferman & Karni, 2010; Smalle, Page, et al., 2017), we used a complex rule set for picture sequences in the visual modality. This allowed us to directly compare long-term visual learning of complex regularities in children of different ages and in adults (studied separately for visuomotor learning: Kóbor et al., 2017 in adults, Tóth-Fáber et al., 2021: 9-15-year-old children), which is important for evaluating whether the previously found age differences are specific to auditory language material or apply to the visual domain as well.

To look into visual sequence learning and transfer across multiple sessions over one year, a task paradigm using an Artificial Grammar (AG) was modified to map within-session and across-session learning trajectories (see Fig. 1): AGL tasks are good at assessing implicit sequence learning in adults (Milne et al., 2018; Pothos, 2007; A. S. Reber, 1967) and children (Nowak & Baggio, 2017; Pavlidou & Williams, 2014; Rosas et al., 2010; Witt & Vinter, 2012). In AGL tasks, participants are first exposed to sequences of elements that follow a certain set of sequence rules (a finite state grammar, see Fig. 2 & *Methods* in Chapter II). After this learning phase, participants are asked to discriminate legal from illegal sequences (test phase) which allows to assess the implicit rule knowledge of participants. Interleaving learning (exposure) with test phases (grammar judgement; Friederici et al., 2013; Milne et al., 2018), allows capturing learning trajectories. Since an AG rule set can be displayed with different items (“surface features”), transfer was tested with a new stimulus set (animals vs. colors).

Participants learned in three sessions on separate days over the course of one week. After one year, first remaining sequence knowledge was tested in three “relearning” sessions with the original item set. In a subsequent session, transfer to a new visual stimulus set was tested that used the same underlying rule set governing item sequences. The learning trajectories and transfer effects of three age groups who took part in the present study (5-year-olds, 6-year-olds, Adults 2) are compared in Chapter III. For two groups (7-year-olds and Adults 1) who served as controls for the other groups after the delay, the study design differed slightly (see Table 1): They both had four instead of three sessions before the delay, including a transfer session in the end. After the 12-month-delay, a subset of these two groups completed the same set of another four sessions at home (as part of an additional follow-up that took part during lockdowns caused by the COVID-19 pandemic). Consequently, these two groups in total performed eight instead of seven sessions of sequence learning, including two transfer sessions (one in the end of each year). Their learning trajectories across one week and one year are discussed separately in Chapter II.

We hypothesized to find an increase in sequence learning performance across sessions in all child groups and adults for the first stimulus set in Year 1. Seven-year-olds, who already completed a subsequent transfer session before the delay, were expected to generalize sequence rules to the second stimulus set at least to the same degree as the adult group (Adults 1).

After the one-year delay, we expected to observe preserved AG knowledge as well as transfer effects at the end of the second set of sessions. Based on an early childhood advantage reported as a higher sensitivity towards sequential regularities, greater forgetting that should promote the extraction of abstract regularities, and stronger overgeneralization in the domains of memory and language, younger age groups were predicted to feature higher retention over one year and quicker relearning of the acquired rule set with the first stimulus set, and to show larger transfer to a second stimulus set in Year 2 (5-year-olds > 6-year-olds > 7-year-olds).

Additionally, we assumed that children predominantly rely on implicit knowledge, while adults acquire more explicit knowledge about the underlying sequence rules. We only covered a restricted age range in children (age 5 to 7 years), making it unlikely to find great differences in reported explicit knowledge between the child groups investigated here. Nevertheless, the expected direction of an age pattern was that older child groups might increasingly acquire more explicit knowledge (5-year-olds < 6-year-olds < 7-year-olds) and possibly depend more on their acquired knowledge for improvement and transfer in sequence learning.

Regarding cognitive skills involved in our sequence learning task, we predicted that better working memory and declarative memory retrieval would benefit AGL task performance. Since our sequence task was tailored to the memory capacities of children age 5-7-years, this association should emerge in children as well (opposed to Smalle, Page, et al., 2017, who used material matched to adult memory capacities). Language grammar skills were hypothesized to be positively correlated with learning outcomes in all age groups, given previous reports that related language processing to sequence learning performance in children and adults.

Chapter II: Visual statistical learning across one year in 7-year-olds and adults (Project 1)[1]

[1] Parts of this chapter are in preparation as a manuscript written by Daniela Schönberger, Patrick Bruns, and Brigitte Röder.

1. Introduction

Both children and adults are capable of tracking sequential information in the environment (Conway, 2020). This ability has been investigated under the terms statistical learning or implicit learning and is referred to as sequence learning throughout this dissertation (see Chapter I). However, it has been proposed that this mechanism is particularly effective in children, allowing them, for example, to quickly and implicitly acquire language (Aslin, 2017; Erickson & Thiessen, 2015; Romberg & Saffran, 2010; Uddén & Männel, 2018). This assumption has mainly been based on cross-sectional evidence which compared how children vs. adults acquire sequential regularities within a single learning session: Firstly, behavioral results suggested a higher sensitivity for visual regularities in young children (< 12 years) compared to older age groups in an implicit learning task (Janacsek et al., 2012). Secondly, neural evidence on implicit learning markers (event-related brain potentials) implied that infants and young children, but not adults, pick up statistical regularities in passive exposure situations, that is, when they are not task-relevant (Friederici et al., 2011; Friederici et al., 2013; Mueller et al., 2018; Rohlf et al., 2017). Both lines of findings were taken as evidence for a shift from more implicit to more explicit learning mechanisms across development (Daltrozzo & Conway, 2014; Janacsek & Nemeth, 2012; Nemeth et al., 2013). This developmental shift might underlie children's superiority in certain aspects of language learning as well, like grammar acquisition or word segmentation. As elaborated in Chapter I, the reliance on implicit vs. explicit learning mechanisms can be experimentally manipulated in adults to counteract the proposed developmental shift towards explicit learning, e.g., by decreasing cognitive control prior to passive exposure to hidden auditory regularities, which fostered implicit word learning (Smalle et al., 2022). Preventing interference from explicit learning mechanisms might consequently help to mitigate the age-related loss in sensitivity towards passively encountered regularities. The view of a high initial sensitivity for statistical regularities which decreases later in childhood has not remained uncriticized. Some authors have even proposed an improvement in implicit learning of sequential and probabilistic relationships from childhood to adulthood (Lukács & Kemény, 2015; Weiermann & Meier, 2012).

After disrupting brain areas that usually contribute to explicit learning by brain stimulation in adults during a first learning phase, delayed learning of visuomotor regularities was reported to be improved when assessed after a 24-hour consolidation period (Ambrus et al., 2020). This suggests that benefits from implicit learning mechanisms might extend to longer timescales and can result in improved performance after a delay period. However, longitudinal designs which directly compare multi-session learning and retention effects between children and adults are largely missing (but see Ferman & Karni, 2010; Smalle, Page, et al., 2017). Thus, it remains unknown whether higher learning capabilities do exist in children and whether they would result in superior memory of and relearning advantages for sequential regularities.

Previous multi-session learning studies with auditory sequences investigated phonological word-form learning (Smalle, Page, et al., 2017) and the learning of an artificial morphological rule (Ferman & Karni, 2010). These studies provided first evidence that a repeated training with a single sequence rule over multiple sessions improves performance in adults and in 8- to 12-year old children alike, and suggested that learning relies more on explicit knowledge of sequence rules in adults than in children. However, these studies yielded inconsistent results regarding the retention of implicitly learned sequences: Smalle, Page, et al. (2017) observed memory advantages in 8 to 9 year-old children for retaining implicitly acquired syllable sequences up to 12 months after their last learning session. For retention after a 12-month delay, this age effect only reached significance for matched subgroups with comparable performance levels before the delay. An explicitly cued sequence was retained equally well by both age groups across 4 hours, 1 week and 12 months (Smalle, Page, et al., 2017). Ferman and Karni (2010) investigated implicit learning of a sequential language rule across several training sessions and observed preserved performance levels and reaction time improvements after a 2 month retention interval for both adults and 12-year-olds, but not for 8-year-olds. Since capabilities for linguistic vs. nonlinguistic stimuli (Raviv & Arnon, 2017; Shufaniya & Arnon, 2018; van der Kant et al., 2020) and adjacent vs. non-adjacent regularities (Uddén & Männel, 2018; Wilson et al., 2020) were reported to follow different developmental trajectories, memory effects for other than auditory stimuli with simple regularities have to be compared between children and adults.

Recent evidence from a delayed learning task in the visuomotor domain has shown that children aged 9 to 15 years (Tóth-Fáber, Janacsek, & Németh, 2021) and young adults (Kóbor et al., 2017) were able to retain both frequency-based (statistical) knowledge and order-based (sequence) knowledge over an interval of 12 months that they had acquired implicitly. Within the group of children, age was not correlated with neither memory for statistical nor for sequence knowledge (Tóth-Fáber, Janacsek, & Németh, 2021). This hint towards age-invariance retention rates was corroborated in a recent study on short-term retention that used the same task in a large cross-sectional sample from the same lab: Nine age groups in the range 7-76 years retained their acquired statistical knowledge to the same degree across 24 hours (Tóth-Fáber et al., 2023; supported by Bayes statistics in favor of no age differences).

Generalization of rule knowledge to new situations is certainly an important skill. In single-session transfer studies, adults and children aged 3-6 years (Nowak & Baggio, 2017) or 6-9 years (Jung et al., 2020) respectively, have been shown to be able to generalize learned regularities to new items in the auditory domain (Nowak & Baggio, 2017) and to new instances of underlying categories in the in the visual domain (Jung et al., 2020). Despite successful transfer in both age groups as reflected in above chance discrimination performance, Jung et al. (2020) reported that adults but not children applied explicit rule knowledge in a transfer situation with high retrieval demands (triplet completion task on category level). Moreover, adults and 12-year-olds, but not 8-year-olds, were observed to generalize a highly practiced language rule to new items in the end of each training session and retained their last performance level for new items in a retention session after a delay of two months. The authors speculated that 8-year-old children lacked transfer effects, because they failed to explicitly discover and report the underlying language rule, which facilitated rule transfer to new items in the older age groups. This interpretation was confirmed in their follow-up study, in which 8 year old children were able to transfer the acquired rule knowledge to new items after they were informed about the rule (Ferman & Karni, 2014). The finding of no (implicit) learning transfer in younger children contradicts predictions derived from other lines of research: There is abundant evidence for a shift from generalization to specificity during development; for example, overregularization of grammatical patterns such as past tense of verb forms, in early stages of language learning is a well-documented finding (Marcus et al., 1992). Others (Keresztes et al., 2018) have used the terms pattern completion (generalization) and pattern separation (specificity) to indicate this trend in memory development. Thus, this literature would predict a stronger rather than lower tendency to generalize learned regularities to new instances early in development.

In the current Project 1, a first long-term perspective on multi-session learning trajectories of children vs. adults is provided on two timescales, across one week and across one year. This serves as a proof-of-concept for the chosen sequence learning task and study design, which was applied in Project 2 as well (see Chapter III). Project 2 then tested how the developmental timing of several sequence learning instances influences learning outcomes in the long run in 5-year-olds, 6-year-olds and adults, while controlling for maturational effects in children after the delay.

Project 1 investigated how children and adults learn *visual* sequences involving *complex rules* over *several sessions* and whether they *generalize* them *to new surface features* in an integrated longitudinal design. This approach extends the existing literature in several ways (further elaborated in Chapter I): (1) Mapping sequence learning across multiple sessions *both* before and after a one-year-delay enabled us to describe learning trajectories of children and adults over an extended period of time. After an initial learning period that comprised four sessions, we investigated relearning of previously acquired regularities one year later again over four equivalent sessions. This approach is different from testing retention in a single follow-up session as typically implemented in previous studies (Ferman & Karni, 2010; Kóbor et al., 2017; Smalle, Page, et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021). (2) We tested the transfer of learning to a new visual stimulus set of a different category in a separate session instead of testing transfer to new items from the trained category (Ferman & Karni, 2010; Jung et al., 2020; Nowak & Baggio, 2017). (3) Instead of a single rule governing syllable sequences in the auditory domain as implemented in previous multi-session studies (Ferman & Karni, 2010; Smalle, Page, et al., 2017), we used a complex rule set governing picture sequences in the visual modality. This allowed us to directly compare long-term visual learning of complex regularities in children and adults (studied separately in both age groups in Kóbor et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021), which helps to evaluating how domain specific vs. domain general reported age differences might be.

An AGL task was implemented, as explained in detail in Chapter I. In AGL tasks, participants are first exposed to sequences of elements that follow a certain set of sequence rules (a finite state grammar, see *Methods*). After this learning phase, participants were asked to discriminate legal from illegal sequences (test phase) which allows assessing implicit rule knowledge. By interleaving learning (exposure) with test phases, we described learning trajectories. Since an AG rule set can be displayed with different items, transfer in the present study was tested with a new set of pictures from different category (animals vs. colors).

The present study recruited adults (Adults 1) and 7-year-old children (see Table 1), because main differences in both sequence learning and transfer were expected for children younger than 8 years as compared to adults. These two age groups are described separately as Project 1 here, since they had the same study design including a transfer session both before and after a long-term delay, serving as control groups for Project 2 with different age cohorts (see Table 1). Characterizing long-term learning trajectories in children vs. adults here is used as a first step for the further investigations in Project 2, as elaborated above.

We hypothesized that both children and adults learn the sequence rules and transfer this knowledge to a new stimulus set. Moreover, we expected to observe preserved AG knowledge after a one year period. Seven-year-olds were expected to quicker implicitly acquire the AG, to show lager transfer to a new stimulus set and to feature higher retention of the acquired rule set over one year. Finally, we predicted that children predominantly rely on implicit knowledge while adults acquire more explicit knowledge about the underlying sequence rules.

2. Methods

2.1. Participants

The study involved 30 healthy children (7 years old ± 2 months at Session 1) from the City of Hamburg, Germany, and 30 healthy young adults, mostly undergraduate students recruited from the University of Hamburg. All participants did not report a history of seeing or hearing impairments nor any neurological disease. They all were native German speakers.

During the course of the study, the data of five participants had to be excluded from the analyses in Year 1. The reasons for exclusion were that it turned out that one adult was not a native speaker of German, data of one child was lost due to technical issues, and three participants did not adhere to task instructions at home (one child: one missing session, one child and one adult: no night between two sessions). The remaining 27 seven-year-olds (15 female, mean age at Session 1: 7.05 ± 0.07 years, range: 6.91-7.20 years) and 28 adults (18 female, mean age at Session 1: 23.12 ± 3.50 years, range: 18.83-33.54 years) were included for the analyses of the first four sessions in Year 1.

From the original sample of 60 participants, 43 participants completed the experiment a second time after approximately 1 year at home[2] (see section *Study design*). The data of three of the returning 43 participants (7-year-olds: n = 22, Adults 1: n = 21) for the home follow-up in Year 2 could not be analyzed due to deviations from task instructions in these sessions (two children: mix-up of stimulus set order, one child: 30 days between two sessions). Since the relearning analysis combined data from Year 1 and Year 2, we additionally had to exclude four of the returning participants since they had missing data in Year 1 for these analyses. This left a total of 16 seven-year-olds (12 female, mean age at first session of Year 2: 8.18 ± 0.09 years, range: 8.03-8.31 years) and 20 Adults 1 (14 female, mean age at first session of Year 2: 23.24 ± 2.48 years, range: 19.94-28.47 years) for the analyses of relearning after the one-year-delay.

All participant characteristics for the final samples analyzed for Year 1 (Table 2) and for joint Analyses of Year 1 and Year 2 (Table 3) are listed below.

Adult participants were compensated with €8 per hour or earned course credit at the end of each session; children received a small toy at the end of each on-site session and after the last session at home in Year 2 in case of a home follow-up. Study-related travel costs

[2] This home follow-up was part of an additional follow-up that took part during lockdowns caused by the COVID-19 pandemic, and had originally not been planned for.

were reimbursed. All participants, i.e., children and adults, and additionally children's legal guardians consented prior to participation (with written consent obtained from adult participants and from children's legal guardians). The study was approved by the Local Ethics Board of the Faculty of Psychology and Human Movement Science at the University of Hamburg, and was conducted in accordance with the ethical guidelines of the Declaration of Helsinki (revised form of 2013).

Table 2

Participant Characteristics of the Final Sample for Analyses of Year 1 across 1 Week

Participant Characteristics	**7-Year-Olds (*n* = 27)**	**Adults 1 (*n* = 28)**
Days between Sessions		
Session 1 to 3	4.11 (1.19)	4.00 (1.59)
Session 1 to Transfer 1	7.00 (0.73)	6.75 (1.08)
Age (years)	7.05 (0.07)	23.12 (3.50)
Age span	6.91 – 7.20	18.83 – 33.54
Gender (f/m)	15/12	18/10
School/Education [a]	*n* = 26 1st grade elementary school *n* = 2 2nd grade elementary school	*n* = 28 university students
Bilinguals	4	6 *
Daily mobile device usage (min) [a]	9.07 (15.45)	254.77 (126.00)
Explicit knowledge Year 1		
N Session 3: 0/1[b]	4/20 (3 NA)	0/28
N Transfer 1: 0/1[b]	7/20	0/28
Score Transfer 1[c]	0.46 (0.23)	0.54 (0.26)

Note. M (SD), betw. = between, NA = missing values for open questions.

* *n* = 25 (3 missing values).

[a] assessed in Year 1 (Session 1).

[b] 0 = sequence rules not mentioned in answers to open questions, 1 = sequence rules mentioned in answers to open questions.

[c] scores could range from -1 (no rule knowledge) to 1 (max. rule knowledge).

Table 3

Participant Characteristics of the Final Sample for joint Analyses of Year 1 & Year 2

Participant Characteristics	**7-Year-Olds (*n* = 16)**	**Adults 1 (*n* = 20)**
Time period betw. Year 1 & 2 (months betw. Session 1 & 4)	13.00 (1.10) [11.00-15.00]	12.75 (0.44) [12.00-13.00]
Days between Sessions of Year 1		
Session 1 to 3	4.19 (0.83)	3.80 (1.54)
Session 1 to Transfer 1	6.88 (0.50)	6.60 (1.14)
Days between Sessions of Year 2		
Session 4 to 6	2.63 (0.81)	2.45 (0.60)
Session 4 to Transfer 2	4.19 (1.52)	3.85 (0.93)
Age Year 1 (years)	7.06 (0.08) [6.92-7.20]	22.14 (2.48) [18.83 – 27.36]
Age Year 2 (years)	8.18 (0.09) [8.03-8.31]	23.24 (2.48) [19.94 – 28.47]
Gender (f/m)	12/4	14/6
Education [a]	*n* = 15 1st grade elementary school *n* = 1 2nd grade elementary school	*n* = 20 university students
Bilinguals	3	6 *
Daily mobile device usage (min) [a]	9.24 (15.46)	250.96 (120.89)
Explicit knowledge Year 1		
N Session 3: 0/1[b]	2/14	0/20
N Transfer 1: 0/1[b]	2/14	0/20
Score Transfer 1[c]	0.51 (0.17)	0.57 (0.26)
Explicit knowledge Year 2		
N Session 6: 0/1[b]	0/16	0/20
N Transfer 2: 0/1[b]	0/16	0/20
Score Transfer 2[c]	0.68 (0.15)	0.66 (0.21)

Note. M (SD) [range], betw. = between, NA = missing values for open questions.

[*] *n* = 17 (3 missing values).

[a] assessed in Year 1 (Session 1).

[b] 0 = sequence rules not mentioned in answers to open questions, 1 = sequence rules mentioned in answers to open questions.

[c] scores could range from -1 (no rule knowledge) to 1 (max. rule knowledge).

2.2. Design and procedure

2.2.1. Study design

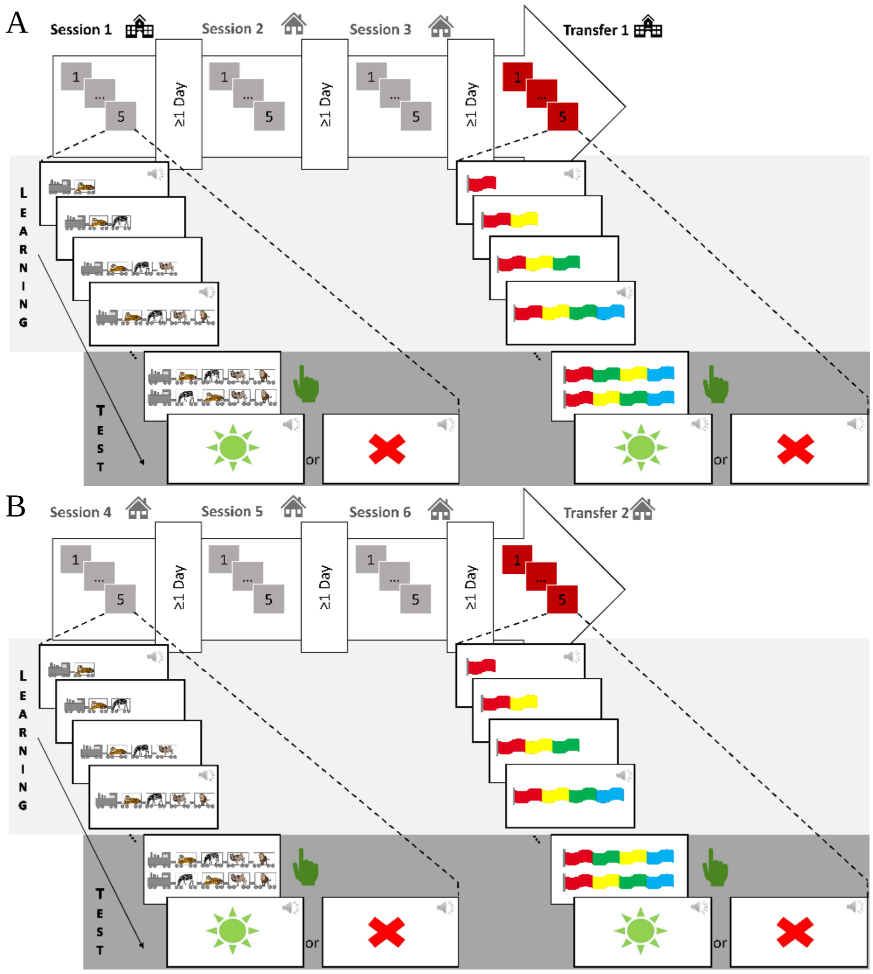
All participants completed a total of four sessions in Year 1 on separate days over the time of approx. one week (see Fig. 1A; mean time span Session 1 to Transfer 1, see Table 2 & 3):

- Session 1 (in the lab): After the assessment of working memory, the first learning session with the tablet computer (with stimulus set 1) followed. Next, we measured declarative memory and German grammar skills (see section *Memory and language skills*). Session 1 lasted 90 to 120 minutes including participant briefing and breaks.
- Session 2-3 (at home): Two more learning sessions took place with stimulus set 1 on the tablet computer (mean time span Session 1 to Session 3, see Table 2 & 3).
- Transfer 1 (in the lab): On the tablet computer stimulus set 2 was introduced with the same underlying rule set to test transfer of AG learning. Moreover, at the end of the session, explicit knowledge about sequence orders was assessed with a questionnaire in adults and adapted questions with picture cards in children (see section *Explicit knowledge of sequence rules*).

Due to the ongoing COVID-19 pandemic, the follow-up in Year 2 took place at home (data sets from n = 16 seven-year-olds and n = 20 young adults were included, see section *Participants*). Thus, for these sessions, participants received all material and a tablet computer by mail and completed four additional learning sessions at home (see Fig. 1B) after approximately one year (mean time span from Session 1 in the final sample: 7-year-olds: see Table 3): Session 4, 5 and 6 used the first stimulus set from Year 1 (that is, the same stimulus set as used in Session 1 to 3), while Transfer 2 employed the second stimulus set from Year 1, that is, the same as used in Transfer 1 (mean time span between Session 4 and Transfer 2, see Table 3). All sessions implemented the same AG rule set. At the end of Transfer 2, explicit sequence knowledge was assessed with the same questionnaires as in Year 1 (Transfer 1) which had been mailed to the parents and participants, respectively. Questions assessing explicit sequence knowledge asked about legal items in salient positions and legal item-item transitions, as detailed in the section *Explicit knowledge of sequence rules*.

Figure 1

Study Design of all 8 Sessions of Visual Sequence Learning in Year 1 (A) & Year 2 (B)



Note. The first 3 sessions of each year used a first stimulus set (here: Stimulus Set Animals), while the last session of each year employed a second stimulus set (here: Stimulus Set Colors) to investigate transfer of learned AG rules. Each session consisted of 5 task blocks with alternating learning (light gray background) and test (dark gray background) phases.

2.2.2. Visual sequence learning task

The present sequence learning task built on previous AGL tasks for children of similar age (esp. Rosas et al., 2010; Witt & Vinter, 2012). In AGL tasks, participants are exposed in a learning phase to sequences of elements which follow a certain set of sequence rules (a finite state grammar, see Fig. 2). In the following test phase, they have to distinguish grammatical sequences from sequences that violate the sequence rules (“ungrammatical” sequences).

We implemented visual sequence learning on a tablet computer with five blocks of alternating learning phases (passive watching) and test phases (two-alternative forced choice task) per session. All instructions were child-directed voice recordings embedded in the task and they were automatically played upon starting the application. This procedure guaranteed a high level of standardization of the sessions at home (see Appendix A for the wording of all instructions).

One task block comprised one learning phase and one subsequent test phase (see Fig. 1). In the learning phases (see Fig. 1 light gray background), participants were instructed to attentively watch 18 grammatical sequences with 3 to 7 items which were randomly taken from the 27 possible grammatical sequences (see section *Construction of grammatical and ungrammatical sequences*). Only one stimulus set, either Stimulus Set Animals with animals in train cars ("circus trains" belonging to a circus director: adapted from Rosas et al., 2010) or Stimulus Set Colors with color segments ("team flags" belonging to a sports team: adapted from Witt & Vinter, 2012) was presented. Which stimulus set served as stimulus set 1 and 2 was counterbalanced across participants.

In the test phases following each learning phase (see Fig. 1 dark gray background), participants were asked in each of the 10 trials to select from the two displayed sequences the sequence they considered as belonging to the previously introduced task character (i.e., the grammatical sequence; two-alternative forced choice). This instruction was repeated for each trial of the test phase and participants selected their choice by touching the selected sequence on the tablet's touchscreen. Each individual test trial comprised two sequences with each being either short (3-5 items, 5 test trials) or long (6-7 items, 5 test trials). Always one of the two sequences was grammatical and one was ungrammatical (see section *Construction of grammatical and ungrammatical sequences*). Participants received audio-visual feedback (see section *Stimuli and apparatus*) after each test trial which indicated if their answer was correct or incorrect.

One session of the sequence learning task (5 blocks with an alternating learning and test phase each) took about 25-30 min to complete; additionally, short breaks were offered to the participants after each block.

2.3. Material

2.3.1. Stimuli and apparatus

Stimuli and timing of the visual sequence learning task. Participants completed the self-programmed visual sequence learning task in the mobile Neurobs Presentation App (Version 3.0.1, Neurobehavioral Systems Inc., 2019) on a 10.1-inch tablet computer (Samsung Galaxy Tab A 10.1).

Grammatical and ungrammatical sequences were 3 to 7 items long and were built from a total of 5 different pictures per stimulus set (Stimulus Set Animals with circus trains: giraffe, camel, lion, tiger and zebra; Stimulus Set Colors with team flags: blue, yellow, green, purple and red; all pictures, see Appendix A). The size of one picture (i.e., one animal in a car or one color segment, respectively) was 240x240 pixels (ca. 4° of visual angle at a viewing distance of 40 cm, which is used in all subsequent calculations of the visual angle). For grammatical sequences, these pictures were assigned to the positions in the artificial grammar (see numbers in Fig. 2) in two different, randomly selected ways per stimulus set to control for position effects of specific animals or colors, respectively (an individual assignment for each participant was technically not feasible due to the implementation in a mobile App). This picture assignment resulted in two task versions available for each stimulus set, which were used in a counterbalanced manner in both age groups (for all four versions with picture assignments, see Table A.1 in Appendix A).

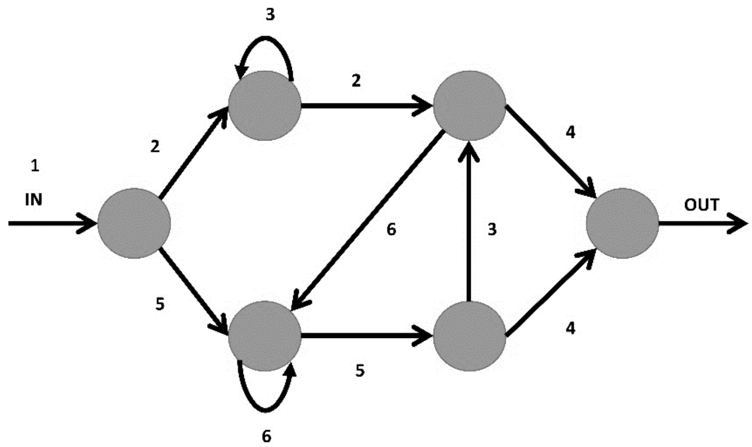
For the animals of the Stimulus Set Animals, we used pictures from the "Multilingual Picture" database (MultiPic; Dunabeitia et al., 2018). The surrounding train (car) features and all other visual stimuli of the task custom made digital drawings. The two chosen stimulus sets (Stimulus Set Animals, Stimulus Set Colors) were adapted from previous studies on sequence learning in children (Rosas et al., 2010 and Witt & Vinter, 2012, respectively).

While the recorded instructions played, the screen showed a picture of a scene corresponding to the stimulus set (Stimulus Set Animals: a circus tent; Stimulus Set Colors: a sports stadium; both retrieved from the MultiPic database by Dunabeitia et al., 2018). To start a learning or test phase, participants were asked to click on a green button in the upper half of the screen (size: visual angle of ca. 4°).

The duration of a single learning trial varied with the number of items of the presented sequence: items, that is single pictures of animal cars or flag segments, built up sequentially (1 s per item), but the full sequence stayed on the screen for 3 s for all sequence lengths. All learning trials were displayed in the middle of the screen. Each trial started and ended with a sound which was presented synchronously with the first and last item of a sequence. Two different sounds, both retrieved from an open-source website (http://www.findsounds.com, accessed 11/16/18; train whistle for circus trains: http://atsf.railfan.net/airhorns/p5.html/, wind blow for team flags: http://www.anzwad.com/dods/sound/ambient), were used, one for each stimulus set. Each learning phase consisted of 18 learning trials amounting to an average duration of about three minutes.

Figure 2

Sequence Rules for the Visual Sequence Learning Task (Artificial Grammar by Reber, 1967)



Note. Grammatical sequences were constructed by following the arrows from “IN” to “OUT” (example sequence: 12324). Instead of numbers, we used pictures of animal cars and color segments (see Fig. 1 & Table A.1) for two stimulus sets that were applied in a counterbalanced manner. Modified drawing of Reber, 1967, Figure 1 (p. 856).

In the test phase all items of the two sequence were simultaneously displayed from the beginning. Each sequence pair stayed on the screen until the participant clicked on one of the two sequences, which in total took on average of about two minutes for all 10 test trials. For the sequence pairs in each test trial, one sequence was displayed in the upper half of the screen, and the other sequence was displayed in the lower half of the screen, with an equal likelihood for the grammatical sequence of appearing in one of the two screen locations.

For audiovisual feedback, a green sun icon with a corresponding sound (correct) or a red cross icon with a corresponding sound (incorrect) was presented in the center of the screen (size of both icons: visual angle of ca. 6°). The two feedback sounds were taken from Leon Guerrero et al. (2016). Participants all started with the same sound volume and were allowed to customize the volume in the beginning of a session, while ensuring that the audio was well audible for each participant.

After the presentation of each sequence (learning phases) and after the response feedback to each trial (test phases), respectively, a gray fixation cross was shown for 1s in the center of the screen (size: visual angle of ca. 0.5°) before the start of the next trial.

Construction of grammatical and ungrammatical sequences. We used the artificial grammar system introduced by Reber (1967) for constructing grammatical sequences (see Fig. 2). Meta-analyses have attested this artificial grammar a relatively low complexity suitable for use in developmental populations (Schiff & Katan, 2014).

From this grammar, 27 grammatical sequences with 3 to 7 items were constructed using the Web App "AGSuite" (Cook et al., 2017; for a full list of all grammatical sequences see Table A.2 in Appendix A), which were divided into 9 "short" sequences consisting of 3 to 5 items and 18 "long" sequences consisting of 6 to 7 items. We then compiled a pool of 140 ungrammatical sequences with 3 to 7 items. For sequences of 5 to 7 items this was achieved by randomly shuffling the middle elements of the grammatical sequences (leaving the most salient first and last item unchanged). For grammatical sequences of only 4 items, either the first 3 or the last 3 items were shuffled. This kept the first or the last item, respectively, unchanged to avoid grammaticality judgements predominantly based on these most salient positions. For the two grammatical sequences of only 3 items, the first item stayed the same, while the second and third item swapped positions (as all other permutations would result in grammatical sequences). Next we computed the "Global Associative Chunk Strength" (ACS) according to Cook et al. (2017, p. 1648) in a slightly modified way for all of the 140 resulting ungrammatical sequences, to match them in difficulty for short (low ACS) and long trials (high ACS): We defined ACS as the sum of an ungrammatical sequence's shared bigrams (item pairs) and trigrams (item triplets) with all bigrams/trigrams existing in all grammatical sequences of the same length (3 to 5 items or 6 to 7 items, respectively), divided by the total number of bigrams/trigrams in the given ungrammatical sequence: A higher ACS means that an ungrammatical sequence shared more picture pairs and triplets with the grammatical sequences of the same length, making this sequence more difficult to be identified as ungrammatical. From the 30 short ungrammatical sequences (3 to 5 items), 11 sequences with a similar ACS (M = 1.44, SD =.18; range: 1.25 to 1.75) were selected, which were later randomly paired with the 9 short grammatical sequences to build the pairs of displayed sequences in the "easy" test trials. Short trials on average had a lower ACS (thus, *easy* trials) than long trials (thus, *difficult* trials), which are described in the following. For the second, "difficult", trial type, we picked 19 out of the 110 long ungrammatical sequences (6 to 7 items) that had a similar ACS (*M* = 5.97, *SD* =.51; range: 5.03 – 6.85) and later paired them randomly with the 18 long grammatical sequences to make up the "difficult" test trials (for a full list of all ungrammatical sequences see Table A.2 in Appendix A). This procedure resulted in 30 ungrammatical sequences (11 short sequences with 3 to 5 items, 19 long sequences with 6 to 7 items) which were presented along with their ACS-matched 27 grammatical sequences (9 short sequences with 3 to 5 items, 18 long sequences with 6 to 7 items) in individual test trials. For each test phase, 5 randomly chosen pairs of "easy" sequences (short sequence with a low ACS) and 5 randomly chosen pairs of "difficult" sequences (long sequence with a high ACS) were presented in a random order, adding up to 10 test trials per block.

Each individual grammatical sequence was restricted to appear at most once per learning and once per test phase, and each individual ungrammatical sequence only once per test phase. Thus, within each task block (comprising 1 learning phase and 1 subsequent test phase), an individual ungrammatical sequence was not seen more than once and an individual grammatical sequence was not encountered more than twice. Therefore, per session any given ungrammatical sequence was presented no more than 5 times and any given grammatical sequence was shown no more than 10 times.

Since random subsets were drawn from the whole set of grammatical sequences for each learning (18 out of 27 sequences) and test phase (5 out of 9 short sequences & 5 out of 18 long sequences), short-term familiarity was an additional dimension of test trials, apart from trial difficulty (based on ACS, see above): a grammatical sequence in a test trial could either have been presented ("seen") or not presented ("not seen") in the directly preceding learning phase of this task block (approx. 2/3 of the grammatical sequences per test phase seen vs. 1/3 not seen).

2.3.2. Explicit knowledge of sequence rules

Explicit knowledge about underlying rules of the sequence learning task was assessed with three open questions at the end of Session 3 and Transfer 1 and a comprehensive questionnaire including additional questions at the end of the final session of *Transfer 1* (see Appendix A) . A shorter version of the questionnaire was administered for children. All assessments were based on a procedure by Whitmarsh et al. (2013).

The same open questions and questionnaires were administered at the end of Session 6 and Transfer 2, respectively. Children's parents were asked to pose the questions and document the answers of their children in Transfer 2, since unlike Transfer 1, this transfer session was run at home. Due to miscommunication, 5 children answered this second questionnaire of Year 2 after Session 6, that is, questions were not asked about stimulus set 2 used for transfer, but instead about stimulus set 1.

The questionnaire asked about sequence knowledge of the just completed sequence learning task and started with three general questions in an open answer format: (1) *What do you think this task was about?*, (2) *Did you have a strategy to choose which train/flag belonged to [introduced person/team]?*, (3) *Did you notice anything about the trains/flags of [introduced person/team]?*. It continued with 11 (adults) or 5 (children) specific questions, respectively, measuring reported knowledge about legal first and last pictures (*With what [animal/color] could the trains/flags of [introduced person/team] begin/end?*) and legal bigram transitions (*What animals/colors could repeat themselves? What animal(s)/color(s) could follow animal/color X? What animal(s)/color(s) could **not** follow animal/color X?*) as described in detail in Whitmarsh et al. (2013). Adults and children assessed at home had to choose from the five possible colored pictures (animals/colors) provided as response options in the questionnaire (multiple choice). Children assessed in the lab chose their response as printed cards (multiple choice), respectively.

The first three questions about the task were considered together as a proxy for overall awareness of sequence rules. For this, we scored if participants mentioned sequence rules in any of their answers (0 = no mention of sequential order in any of the three questions [(1),(2),(3)]; 1 = mention of sequential order in at least one of the three questions).

From the specific questions, we calculated an explicit knowledge score for Year 1 and Year 2, respectively. To this end, correct answers (correctly chosen pictures) were added up and incorrect answers (incorrectly chosen pictures) were subtracted from the correct answers in a weighted manner (see below), then the result was divided by the number of questions answered. We modified the assessment and thus the scoring from Whitmarsh et al. (2013), to render scores comparable for both age groups: On order to make the assessment feasible for children, we randomly picked two (out of five) pictures and asked about their bigram transitions, instead of asking about all bigram transitions as in adults. The answer to each question in both age groups was hence weighted according to the probability of valid bigram answers (i.e., the number of correct pictures chosen by a participant was divided by the number of all possible correct pictures [considering misses] and in an analogue manner for incorrect picture choices [considering correct rejections]). This was done because the number of valid bigram answers differed for different pictures (as they represented different “arrows” in the artificial grammar, see numbers in Fig. 2). As children were not asked about bigram transitions for all pictures, but only about those of two out of five pictures picked randomly, this weighting was used to render answers comparable across participants. This means that for each question, a score from -1 (no correct answers given) to 1 (all correct answers given, without any false alarms) could be obtained. The resulting scores for each question were then added up and divided by the number of questions answered in total, to obtain the final score for explicit sequence knowledge (range of total score: -1 to 1). Higher scores indicated more explicit sequence knowledge.

2.3.3. Memory and language skills

To assess working memory, declarative memory and German grammar skills, we administered equivalent psychometric tests in Session 1 (Year 1) in both age groups and normalized all test scores according to age (except for Plural German Grammar Skills in Adults 1 for which norms were not available and for which hence raw scores were analyzed).

Descriptive data for both age groups in the assessed memory skills and grammar skills, are detailed in Chapter IV. There, correlational analyses relate AGL performance to cognitive skills in all age groups investigated in this dissertation.

2.3.4. Additional assessments

Additional information about children and adults was collected with custom-made questionnaires at the end of Session 1, assessing their visual and hearing development, educational background, 2nd languages and the use of (mobile) devices. A screening tool for behavior on clinically relevant dimensions (Child Behavior Checklist CBCL, Döpfner et al.,2014; Brief Symptom Checklist BSCL, Franke, 2017) was administered in this context as well. Adult participants filled out these questionnaires themselves, while caregivers did so for participating children.

2.4. Data analysis

We characterized learning trajectories and averaged performance scores as proportion of correct test trials: Across-session learning was assessed as the mean performance of 50 test trials of each session. Within session learning trajectories were derived based on the means of 10 test trials per block.

Trials with shorter reaction times than 200 ms were disregarded, since we did not consider it feasible to successfully process the two sequences within less than this time. This exclusion criterion reduced trial numbers by 0.18 % (a total of 20 trials from 5 seven-year-olds were excluded) for analyses of Year 1 (see results section *Repeated learning across one* *week (Year 1)*) and by 0.12 % (a total of 17 trials from 7 seven-year-olds and from 1 adult were disregarded) of Year 1 and Year 2 combined for relearning analyses (see results section *Relearning after a one-year-delay (Year 1 vs. Year 2)*).

To compare performance changes over time between the two age groups, repeated-measures Analyses of Variance (ANOVAs) were conducted using the *ez* package in R (Lawrence, 2016), with *Age* (7-year-olds vs. Adults 1) as between-subject factor and *Session* (levels depending on analyses as described below and in the respective *Results* section) as within-subject factor.

1. Within Year 1, the following comparisons were analyzed:
    - Session 3 vs. Session 1 (Learning Gains)
    - Transfer 1 vs. Session 1 (Transfer Savings)
    - Transfer 1 vs. Session 3 (Transfer Loss)
2. For comparing start and end levels between Year 1 and Year 2, the following sessions were analyzed:
    - Session 4 vs. Session 1 (Start Level)
    - Session 4 vs. Session 3 (Retention)
    - Session 6 vs. Session 3 (End Level)
3. For comparing session differences between Year 1 and Year 2 (performance improvement over 3 sessions, transfer performance relative to the first and last session with the first stimulus material), an additional within-subject factor *Year* (Year 1 vs. Year 2) was added in the ANOVAs, resulting in the within-subject factors *Session* and *Year* with levels described in the respective *Results* section.

ANOVAs were followed up with appropriate post-hoc tests. If scores were not normally distributed, Wilcoxon signed rank tests (instead of *t*-tests) and Spearman correlation coefficients (r_s) were calculated. Two tailed significant (<.05) *P*-values (if not indicated otherwise) were Greenhouse-Geisser-corrected (in case of violated sphericity) or Holm-corrected (in case of multiple comparisons). Effect sizes were calculated as generalized eta squared (η^2_g) for ANOVAs, as Cohen's *d* for *t*-tests and as matched rank biserial correlation (*r*) for Wilcoxon singed rank tests, respectively.

For session comparisons that involved Session 1 or Session 4, additional control analyses were conducted with proportion correct of test trials averaged over block 2 to 5 of each session (without the first block), to account for task novelty. The rationale was to control for differences in general task familiarity between sessions, since performance in the very beginning might be poorer compared to later sessions because of a lower familiarity with the general task setting.

Due to the home-setting, there were some additional deviations from the task instructions in the final sample, mainly in Year 2 (3 participants with 4 instead of 3 sessions with the first stimulus set, 2 participants with 6 instead of 5 task blocks in 1 session, 1 participant with 4 instead of 5 task blocks in Session 5). These participants were included in the final analyses, since they did not show any systematic peculiarities in their response patterns and only the originally scheduled trials were included in case of additional task blocks completed (*n* = 5) or session performance was averaged over the available data in case of a missing task block (*n* = 1), respectively. Additionally, we checked for each of the reported analysis whether excluding these six participants with slightly different task exposure would qualitatively change the pattern of results.

We performed equivalent Bayesian analyses for all inferential statistical analyses in the software JASP (Version 0.14.1; JASP Team, 2021), using default priors, and report the Bayes Factor (BF_{10}). The *BF* helps to evaluate whether the data at hand support the null-hypothesis (H_0) or the alternative hypothesis (H_1), and has been described as a suitable tool for interpreting null results (Dienes, 2014). For main and interaction effects in ANOVAs, we report the inclusion Bayes factor (BF_{incl}) – which compares models that contain the effect of interest to equivalent models stripped of this effect – as implemented in JASP Version 0.14.1 and recommended by e.g. Mathôt (2017) and Quintana & Williams (2018). *BF* values between 1/3 and 1/10 indicate moderate evidence for the H_0, while a BF of lower than 1/10 is considered strong evidence for the H_0; a BF between 1 and 1/3 is defined as anecdotal evidence for the H_0 (Schönbrodt & Wagenmakers, 2018). On the other hand, *BF* values between 3 and 10 indicate moderate evidence for the H_1, while a *BF* from 10 onwards is considered as strong evidence for the H_1 and a *BF* between 1 and 3 is defined as anecdotal evidence for the H_1 (Schönbrodt & Wagenmakers, 2018). For post-hoc tests on scores that were not normally distributed, the *BF* was calculated for non-parametric test equivalents to the respective inferential tests and is reported using the default setting of data augmentation algorithms with 5 chains of 1000 iterations as implemented in JASP.

All data analyses apart from Bayesian analyses were performed in the software R (Version 4.1.0; R Core Team, 2021).

To further look into within-session learning, we made use of the trail-by-trial response data and fit the state-space random effects model by Smith et al. (2005) to binary responses (correct = 1, incorrect = 0) in all 150 test trials of the 3 sessions with stimulus set 1, separately for each age group and each year (Session 1 to 3 in Year 1: data from 27 children and 28 Adults 1; Session 4 to 6 in Year 2: data from 15 children and 20 Adults 1). This model estimated the first trial as the timepoint where learning had first occurred for the whole population (i.e., age group), by estimating an unobservable learning state process, defined as a random walk. For an estimation of the learning curves it used a state-space random effects model and Expectation-Maximization algorithm, characterizing the dynamics of the learning process as a function of trial number (Smith et al., 2005). The modeling script was provided in Matlab (Matlab, MathWorks 2020) from the website indicated by Smith et al. (2005; http://annecsmith.net/behaviorallearning.html). The estimated first learning trial from this population modeling was then used to compare within-session learning between age groups.

3. Results

3.1. Repeated learning across one week (Year 1)

We first tested whether both age groups performed above chance in every session, to assess whether they had learned the AG rule (Fig. 2A). For all sessions, group-level performance exceeded the chance level of 0.5 (two-alternative forced-choice trials) in both 7-year-olds (all $t(26) \geq 2.80$, all $p \leq .005$, all $d \geq 0.54$, all $BF_{+0} \geq 9.53$; one-sided) and in Adults 1 (all $V(27) = 406.00$, all $p \leq .004$, all $r \geq 0.87$, all $BF_{+0} > 100$; one-sided). For session means and additional information on both groups, see Table 4A and Appendix B (session averages without block 1). All following analyses are based group *Means(SDs)* in Table 4A (n = 27 7-year-olds & n = 28 Adults 1) which will be referenced for better readability.

Table 4

A *Year 1: Proportion Correct in AGL per Session and Age Group*

	Session 1		**Session 2**		**Session 3**		**Transfer 1**	
	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1
N	27	28	27	28	27	28	27	28
M	.55	.78	.62	.90	.65	.93	.63	.89
SD	.09	.09	.12	.07	.16	.07	.11	.11
Min	.30	.62	.42	.76	.35	.74	.36	.60
Max	.70	.98	.88	1.00	.96	1.00	.90	1.00
*N**	16	20	16	20	16	20	16	20
M	.56	.79	.69	.92	.69	.94	.67	.90
SD	.07	.10	.13	.06	.17	.07	.09	.12
Min	.44	.62	.50	.76	.46	.74	.52	.60
Max	.70	.98	.88	1.00	.96	1.00	.90	1.00

B *Year 2: Proportion Correct in AGL per Session and Age Group*

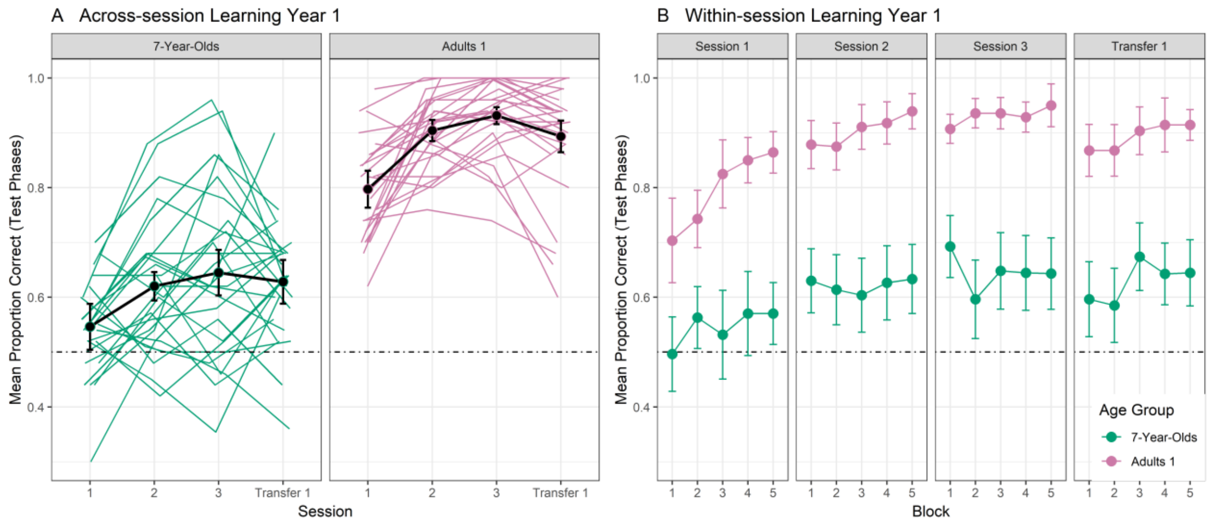
	Session 4		**Session 5**		**Session 6**		**Transfer 2**	
	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1
N	16	20	16	20	16	20	16	20
M	.71	.91	.76	.94	.80	.94	.73	.92
SD	.11	.10	.16	.09	.13	.09	.13	.09
Min	.52	.60	.48	.70	.52	.68	.49	.64
Max	.90	1.00	.98	1.00	1.00	1.00	.98	1.00

Note. 7yo = 7-year-olds, Ad 1 = Adults 1, Min = minimal value, Max = maximal value

* Subgroup of returning participants with data for Year 1 & Year 2 (see section *Participants*).

Figure 3

Performance Trajectories Across Sessions and Within Sessions of Year 1



Note. A: Mean proportion of correct responses in the test phases of Session 1 to 3 and Transfer 1 (Year 1) for 7-year-olds (left) and adults (right). Learning curves of single participants are depicted in green (7-year-olds) and pink (adults). B: Mean proportion of correct responses in the test phases of each block of Session 1 to 3 and Transfer 1 (Year 1) for 7-year-olds (green) and adults (pink). The dotted horizontal lines mark chance level performance. Error bars indicate 95% CIs corrected for within-subject comparison according to Morey (2008).

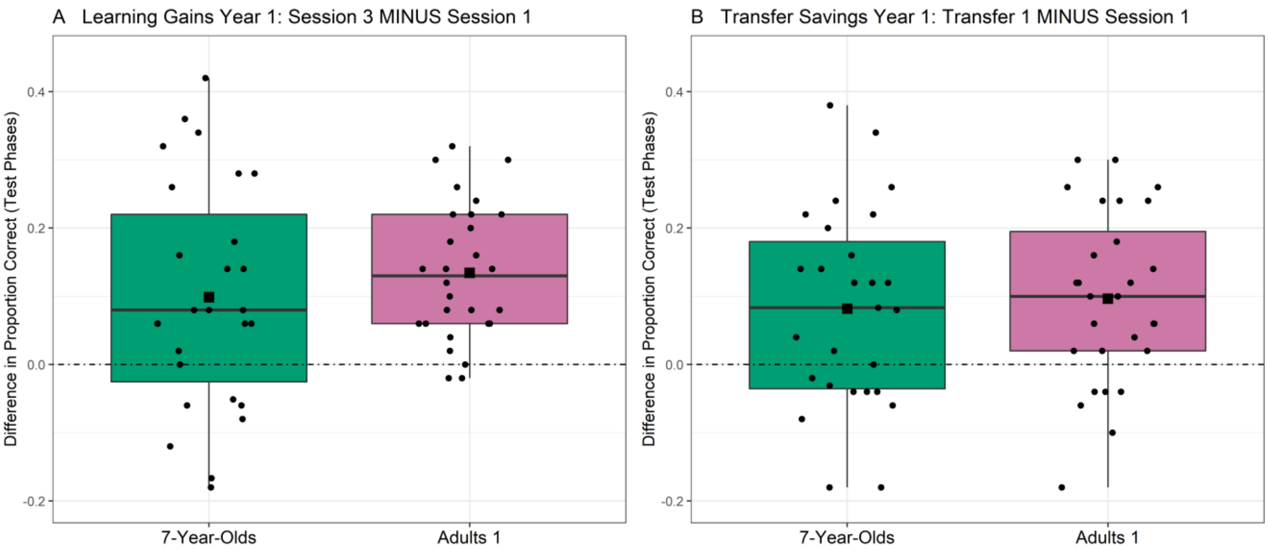
3.1.1. Age comparison: Performance improvement and transfer to new stimuli

To evaluate benefits of repeated sequence training, we tested to what degree performance improved from the first (Session 1) to the last session (Session 3) with the first stimulus set. To this end, averaged session scores (proportion correct of 50 test trials) were entered into an ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 1, Session 3): Learning performance improved in both groups from the Session 1 to Session 3 (see Table 4; main effect of *Session:* $F(1, 53) = 40.59$, $p < .001$, $\eta^2_g = .24$; $BF_{incl} > 100$), as reflected in the positive mean difference scores depicted in Figure 4A. The mean proportion of correct responses was higher in Adults 1 compared to 7-year-olds (see Table 4) for both sessions (main effect of *Age:* $F(1, 53) = 152.93$, $p < .001$, $\eta^2_g = .63$, $BF_{incl} > 100$), while Learning Gains from Session 1 to Session 3 did not significantly differ between groups (interaction *Age*Session*: $F(1, 53) = 0.95$, $p = .333$, $\eta^2_g = .01$; $BF_{incl} = 0.40$).

The same pattern of results emerged when the first block of Session 1 was excluded from the analysis to avoid trivial task familiarity effects (significant main effects of *Age* and *Session:* both $F(1, 53) \geq 22.78$, both $p < .001$, both $\eta^2_g \geq .15$, both $BF_{incl} > 100$, but no significant interaction *Age***Session*: $F(1, 53) = 1.14$, $p = .290$, $\eta^2_g < .01$, $BF_{incl} = .41$).

Figure 4

Learning Gains and Transfer Savings in Year 1



Note. Learning Gains were defined as the difference in the proportion correct in Session 3 minus the proportion correct in Session 1 (A). Transfer Savings were computed as the difference in proportion correct in Transfer 1 minus the proportion correct in Session 1 (B). Boxplots for 7-year-olds (green) and adults (pink) with the groups' median indicated by a black line and the corresponding mean indicated by a black square. Black dots represent single-subject data. Scores above the dotted lines indicate successful learning.

A polynomial contrast analyses separately run for each group provided evidence for a linear increase of performance from Session 1 over Session 2 to Session 3 in 7-year-olds ($p < .001$; see trajectories Session 1 to 3 in Fig. 2A left panel), and for both a linear ($p < .001$) and a quadratic trend ($p = .006$) over these three sessions in the adult data (Fig. 2A right panel). In fact, Adults 1 had reached ceiling by the end of Session 2.

Next, we asked whether children and adults were able to apply their acquired rule knowledge of the first three sessions to the second stimulus set in the transfer session (Transfer 1; see Fig. 4B). To examine transfer savings to the new stimulus set, averaged session scores were entered into an ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 1, Transfer 1). Adults outperformed children (see Table 4), irrespective of the session (main effect of *Age*: *F*(1,

53) = 180.35, $p < .001$, $\eta^2_g = .64$; $BF_{incl} > 100$). Crucially, in both groups performance in the transfer session exceeded performance in the first session (*Mean(SD)* for both groups see Table 4; main effect of *Session:* $F(1, 53) = 23.70$, $p < .001$, $\eta^2_g = .18$; $BF_{incl} > 100$), indicating successful transfer to the new stimulus set (see Fig. 4B). This transfer effect did not significantly differ in size between adults and 7-year-olds (interaction *Age*Session*: $F(1, 53) = 0.16$, $p = .694$, $\eta^2_g < .01$; $BF_{incl} = 0.35$).

A control analysis without block 1 yielded similar results (significant main effects of *Age* and *Session:* both $F(1, 53) \geq 17.06$, both $p < .001$, both $\eta^2_g \geq .12$, both $BF_{incl} > 100$, but no significant interaction *Age*Session*: $F(1, 53) < 0.01$, $p = .934$, $\eta^2_g < .01$, $BF_{incl} = .26$).

We further tested whether performance changed from the last session with the first stimulus set (Session 3) to the subsequent transfer session (Transfer 1): An ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 3, Transfer 1) revealed a significant main effect of *Age* ($F(1, 53) = 106.37$, $p < .001$, $\eta^2_g = .60$; $BF_{incl} > 100$) with adults performing better than 7-year-olds across all sessions (see Table 4). No significant performance loss (main effect of *Session*: $F(1, 53) = 3.08$, $p = .085$, $\eta^2_g = .01$; $BF_{incl} = .74$) or age difference in their preserved performance level in the transfer session (interaction *Age*Session* $F(1, 53) = 0.47$, $p = .498$, $\eta^2_g < .01$; $BF_{incl} = .31$) was obtained.

Taken together, across one week in Year 1, both children and adults learned the AG and finally successfully applied this knowledge to a new stimulus set. Throughout the sessions, adults performed at a higher level than children, while leaning gains did not differ between groups.

3.1.2. Identifying the first learning trial from modeling trial-by-trial performance

For within-session learning, the state-space model by Smith et al. (2005) identified the very first test trial (i.e., in the beginning of Block 1 in Session 1) in the adult group as the first timepoint at which learning had taken place. This means that adults showed within-session learning effects after being exposed to a single learning phase of 18 grammatical sequences. In contrast, for children, the 30th test trial (i.e., at the end of Block 3 in the second half of Session 1) was identified as the first timepoint when learning had happened. Thus, children needed more learning trials until they featured successful learning of the AG in Year 1.

3.1.3. Performance correlations with explicit sequence knowledge

After *Session 3*, 83% of 7-year-olds, as opposed to all adults, spontaneously mentioned to have noticed some aspect related to sequence rules when asked about the AGL task and their decision strategies for stimulus set 1 (open questions, see *Explicit Knowledge of Sequence Rules;* for answers per age group see Table 2). After *Transfer 1*, 74% of 7-year-olds reported having become aware of sequence rules for stimulus set 2, while all adults reported to have done so.

Separate χ^2 -tests for these two timepoints yielded no statistically significant results for comparing proportions of aware vs. unaware participants between both age groups in Session 3 ($\chi^2(1) = 1.58$, $p = .209$, $BF_{10} = .20$) and a marginally significant result for Transfer 1 ($\chi^2(1) = 3.89$, $p = .049$, $BF_{10} = .66$), with more adults than children reporting awareness.

When comparing specific sequence knowledge (assessed at the end of Transfer 1) between the two age groups, scores of reported explicit knowledge did not significantly differ between 7-year-olds and adults ($t(53) = -1.17$, $p = .249$, $d = 0.31$, $BF_{10} = .48$). In 7-year-olds, explicit knowledge about sequence rules (assessed at the end of Transfer 1) was marginally positively correlated with Transfer Savings (Transfer 1 – Session 1; $r_s = .42$, $p = .058$, $BF_{10} = 1.68$; adults n.s.: $r_{s\,=}$ .25; $p = .193$, $BF_{10} = .55$). All other correlations of explicit sequence knowledge with AGL Learning Gains and Transfer Effects are detailed in Appendix B (7-year-olds: all $r_s \leq .33$; $p \geq .176$, $BF_{10} \leq 1.11$; Adults 1: all $r_s \leq .24$, $p \geq .448$, $BF_{10} \leq .42$).

3.1.4. Effects of trial difficulty and familiarity on task performance

Since test trials in the sequence learning task differed with regard to difficulty (short trials with low ACS vs. long trials with high ACS) and short-term familiarity (grammatical sequence of a test trial seen vs. not seen in the directly preceding learning phase; for both trial properties, see *Construction of Grammatical and Ungrammatical Sequences*), we evaluated how these task characteristics related to learning performance averaged across all four sessions of Year 1 in both age groups.

Two different ANOVA models were calculated. The first included difficulty as an independent variable (factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Trial Type* (within-subject; levels: simple vs. difficult seen vs. not seen for short-term familiarity), the other short-term familiarity (factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Trial Type* (within-subject; levels: seen vs. not seen). They both revealed main effects of *Age* (both $F(1, 53) \geq 165.43$, $p < .001$, $\eta^2_g \geq .71$, $BF_{incl} > 100$) and *Trial Type* (both $F(1, 3) = 8.29$, $p \leq .006$, $\eta^2_g \geq .03$ $BF_{incl} \geq 6.79$), but no significant *Age*Trial Type* interaction (both $F(1, 53) \leq 1.33$, $p \geq .255$, $\eta^2_g < .01$, both $BF_{incl} \leq .48$): Overall, adults outperformed children in all trial types and participants performed worse in more challenging trial types (long test trials with high ACS and test trials with a grammatical sequence not seen in the previous learning phase, respectively).

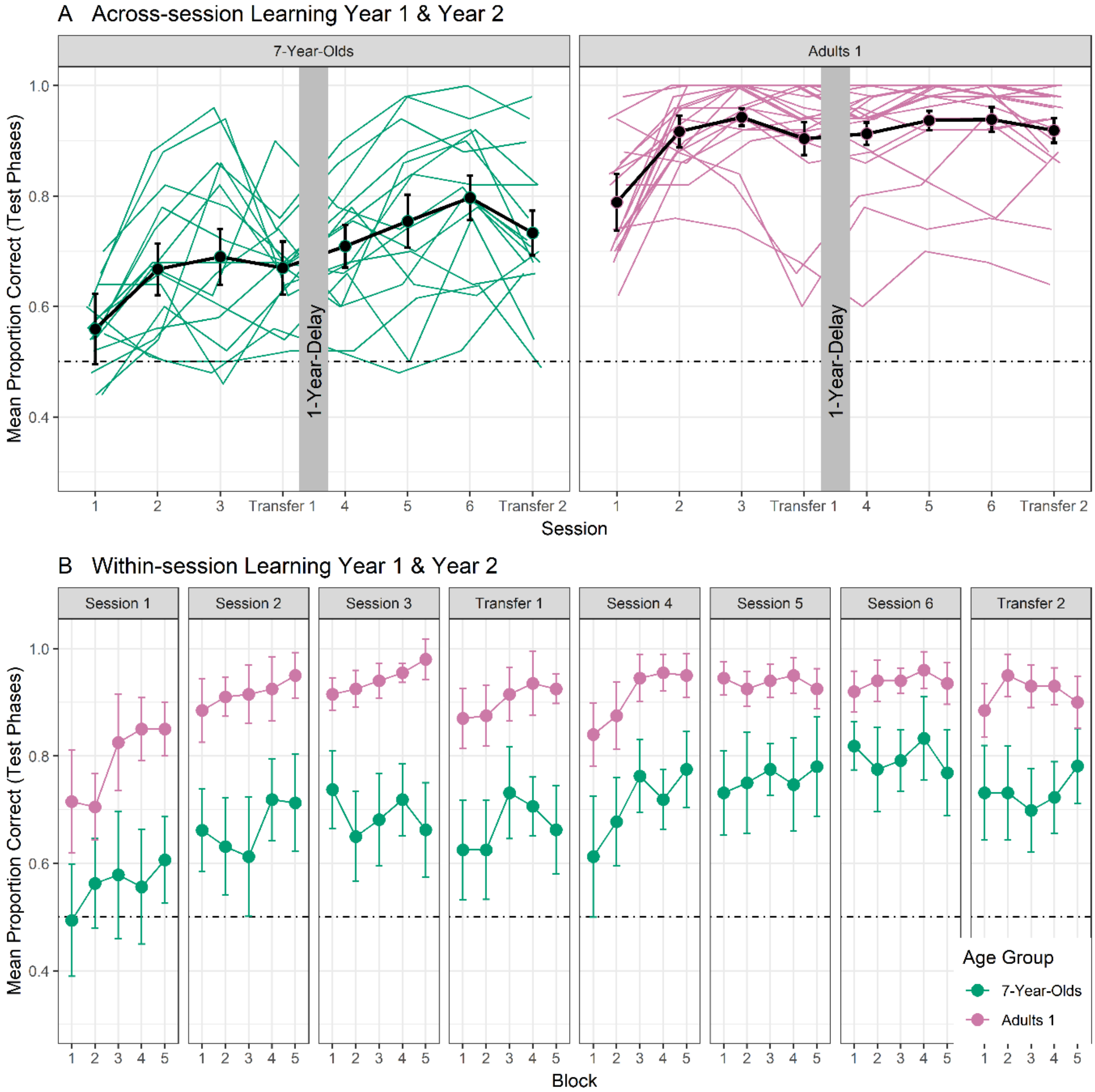
Despite being outperformed by the adult group, the group of 7-year-olds performed on the more challenging types of sequences better than chance (long trials with high ACS: $t(26) = 6.25$, $p < .001$, $d = 1.02$, $BF_{10} > 100$; trials with a grammatical sequence not seen in the preceding learning phase: $t(26) = 5.00$, $p < .001$, $d = .94$, $BF_{10} > 100$).

3.2. Relearning after a one-year delay (Year 1 vs. Year 2)

The following analyses were performed for the subgroup of 16 seven-year-olds and 20 adults (Adults 1), who completed the four home follow-up sessions in Year 2 in addition to all sessions in Year 1. Thus, we first tested, whether this subgroup performed above chance in both Year 1 and Year 2 (see black group means in Fig. 5). For all eight sessions, group-level performance exceeded the chance level of 0.5 (two-alternative forced-choice trials) in 7-year-olds (all $t(15) \geq 3.30$, all $p \leq .008$, all $d \geq 0.83$, all $BF_{+0} > 100$; one-sided) and in adults (one-sided; all $V(19) = 210.00$, all $p \leq .008$, all $r = 0.88$, all $BF_{+0} > 100$; one-sided). For session means and additional descriptive information for both age groups that all following analyses are based on, see Table 4B and Appendix B (session averages without block 1). Table 4 will be referenced for better readability in all following analyses (Table 4A: see data for $n = 16$ seven-year-olds & $n = 20$ Adults 1), instead of providing groups *Means(SDs)* individually for each analyses.

Figure 5

Performance Trajectories Across Sessions and Within Sessions of Year 1 & Year 2



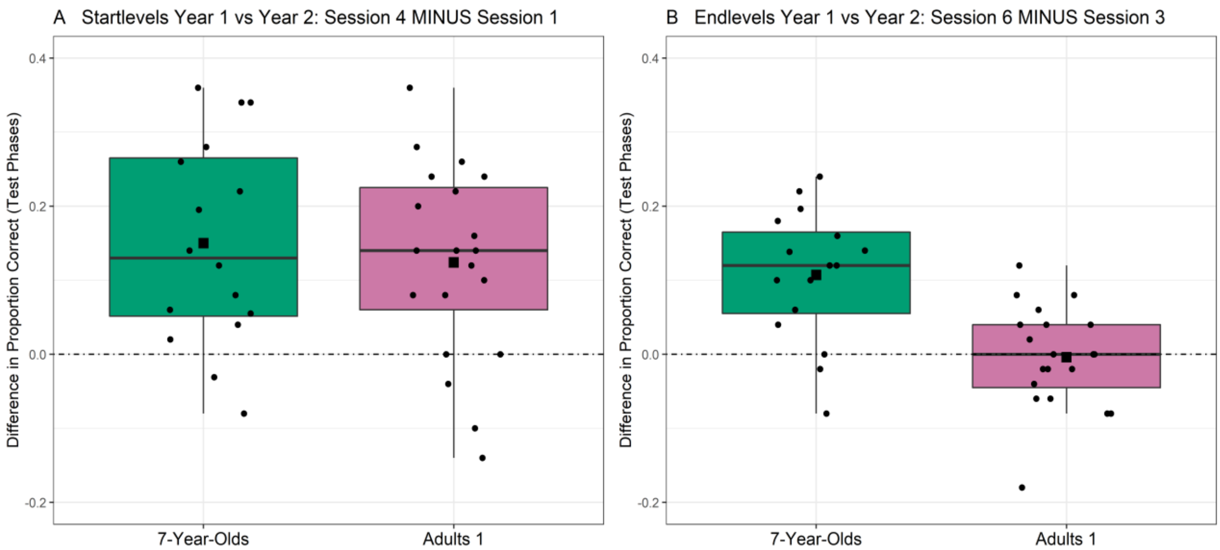
Note. A: Mean proportion of correct responses in the test phases of each session for 7-year-olds (left) and adults (right). Learning curves of single participants are depicted in green (7-year-olds) and pink (adults). B: Mean proportion of correct responses in the test phases of each block for 7-year-olds (green) and adults (pink). The dotted horizontal lines mark chance level performance. Error bars indicate 95% CIs corrected for within-subject comparison according to Morey (2008).

3.2.1. Initial and final performance levels in Year 1 vs. Year 2

To evaluate gains for relearning in Year 2 from the previous learning experience in Year 1, we tested averaged performance in the very first session (Session 1) against averaged performance in the first session of relearning in Year 2 (Session 4) (see Fig. 6A): An ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 1, Session 4) revealed higher overall performance levels in adults than in children (see Table 4; main effect of *Age* $F(1, 34) = 88.02$, $p < .001$, $\eta^2_g = .57$; $BF_{incl} > 100$). Importantly, both age groups performed better in the first relearning session of Year 2 than in the first session of Year 1 (see Table 4; main effect of *Session* $F(1, 34) = 36.92$, $p < .001$, $\eta^2_g = .35$; $BF_{incl} > 100$), as indicated by the positive mean difference scores shown in Figure 7A. The performance gain at relearning in Year 2 was of similar magnitude for 7-year-olds and adults (Age*Session interaction $F(1, 34) = 0.33$, $p = .569$, $\eta^2_g < .001$; $BF_{incl} = .35$).

Figure 6

Session Differences of Year 1 & Year 2 for Initial and Final Performance Levels



Note. Initial performance was compared as difference in proportion correct in the test phases of Session 4 minus Session 1 (A). Final performance was compared as difference in proportion correct in the test phases of Session 6 minus Session 3 (B). Boxplots for 7-year-olds (green) and adults (pink) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

An additional ANOVA, comparing only performance in blocks 2 to 5 for each session, revealed the same pattern of results (significant main effects of *Age* and *Session:* both $F(1, 34) \geq 29.17$, both $p < .001$, both $\eta^2_g \geq .27$, both $BF_{incl} > 100$, and no significant interaction of *Age***Session*: $F(1, 34) = .20$, $p = .532$, $\eta^2_{g\,=}\ .01$, $BF_{incl\,=}\ .29$).

Next, we tested whether both age groups had lost in performance from the last session with the first stimulus set (Session 3) to the first session of relearning of Year 2 (Session 4). An ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 3, Session 4) revealed only a significant main effect of *Age* ($F(1, 34) = 42.81$, $p < .001$, $\eta^2_g = .52$; $BF_{incl} > 100$), with adults performing better than 7-year-olds, irrespective of the session (see Table 4). Participants retained their performance level from the end of Year 1 in Year 2 (see Table 4; main effect of *Session*: $F(1, 34) = 0.11$, $p = .740$, $\eta^2_g < .01$; $BF_{incl} = .25$), with no significant age difference in adults' and children's retention (interaction *Age*Session* $F(1, 34) = 2.43$, $p = .128$, $\eta^2_g = .01$, $BF_{incl} = 0.81$). Excluding block 1 from both sessions' average scores, yielded the same results for the main effects of *Age* ($F(1, 34) = 41.52$, $p < .001$, $\eta^2_g = .51$, $BF_{incl} > 100$) and *Session* ($F(1, 34) = 1.51$, $p = .227$, $\eta^2_g < .01$, $BF_{incl\,=}\ .31$), but rendered the age difference in adults' and children's retention statistically significant (*Age*Session:* $F(1, 34) = 6.18$, $p = .018$, $\eta^2_g = .03$, $BF_{incl\,=}\ 2.96$): Children performed better after the one-year-delay in Session 4 than in the end of Year 1 in Session 3 ($t(15) = -1.89$, $p = .158$, $d = 0.47$, $BF_{10\,=}\ 1.07$), while adults did not (*Mean(SD)* for both groups see Table B.1 in Appendix B; $t(19) = 1.46$, $p = .160$, $d = 0.33$, $BF_{10\,=}\ .58$).

We next examined whether relearning in Year 2 results in a higher final level than in Year 1. Averaged performance in the last session with the first stimulus material was compared between Session 3 and Session 6 (see difference scores for both age groups in Fig. 6B). An ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1) and *Session* (within-subject; levels: Session 3, Session 6) yielded significant main effects of *Age* ($F(1, 34) = 28.17$, $p < .001$, $\eta^2_g = .42$; $BF_{incl} > 100$) and *Session* ($F(1, 34) = 15.37$, $p < .001$, $\eta^2_g = .05$; $BF_{incl} = 5.29$), as well as a significant interaction of *Age*Session* ($F(1, 34) = 17.85$, $p < .001$, $\eta^2_g = .06$; $BF_{incl} = 83.03$): Only 7-year-olds performed better in Session 6 of Year 2, compared to Session 3 of Year 1 (*Mean(SD)* for both groups see Table 4; $t(15) = -4.83$, $p < .002$, $d = 1.21$, $BF_{10} > 100$; shown in red in Fig. 6B left panel), while adults' final performance levels did not differ for Session 3 of Year 1 and Session 6 of Year 2 (*Mean(SD)* for both groups see Table 4; $t(19) = 0.26$, $p = .799$, $d = 0.06$, $BF_{10} = .24$; shown in blue in Fig. 6B right panel).

3.2.2. Performance improvement and transfer effects in Year 1 vs. Year 2

To compare performance improvements over three sessions with the first stimulus set in both years, we tested initial learning in Year 1 (see Fig. 7A) against relearning in Year 2 (see Fig. 7B): An ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1),*Year* (within-subject; levels: Year 1, Year 2) and *Session* (within-subject; levels: First Session [Year 1: Session 1; Year 2: Session 4], Last Session[Year 1: Session 3; Year 2: Session 6]) revealed significant two-way interactions of *Age*Year* ($F(1, 34) = 7.83$, $p = .008$, $\eta^2_g = .03$; $BF_{incl} = 3.26$) and *Year*Session* ($F(1, 34) = 12.79$, $p = .001$, $\eta^2_g = .04$; $BF_{incl} = 23.13$) as well as main effects of *Age, Year* and *Session* (all $F(1, 34) \geq 25.46$, all $p < .001$, all $\eta^2_g \geq .11$, all $BF_{incl} > 100$; all other $F(1, 34) \leq 3.16$, $p \geq .085$, $\eta^2_g \leq .01$, $BF_{incl} < .69$): Overall, participants improved to a greater degree over three sessions in Year 1 (Session 1 vs. Session 3, see Fig. 7A) than they did in Year 2 (Session 4 vs. Session 6, see Fig. 7B; performance improvement in Year 2 [Session 6 – Session 4] vs. performance improvement in Year 1 [Sessions 3 - 1]: $t(35) = 3.63$, $p < .002$, $d = 0.60$, $BF_{10} = 34.34$), and 7-year-olds in general gained more from learning in Year 2 than adults (performance difference average Year 2 [Sessions 4 & 6] – average Year 1 [Sessions 1 & 3] in children vs. Adults 1: $t(34) = 2.47$, $p = .019$, $d = 0.83$, $BF_{10} = 3.12$).

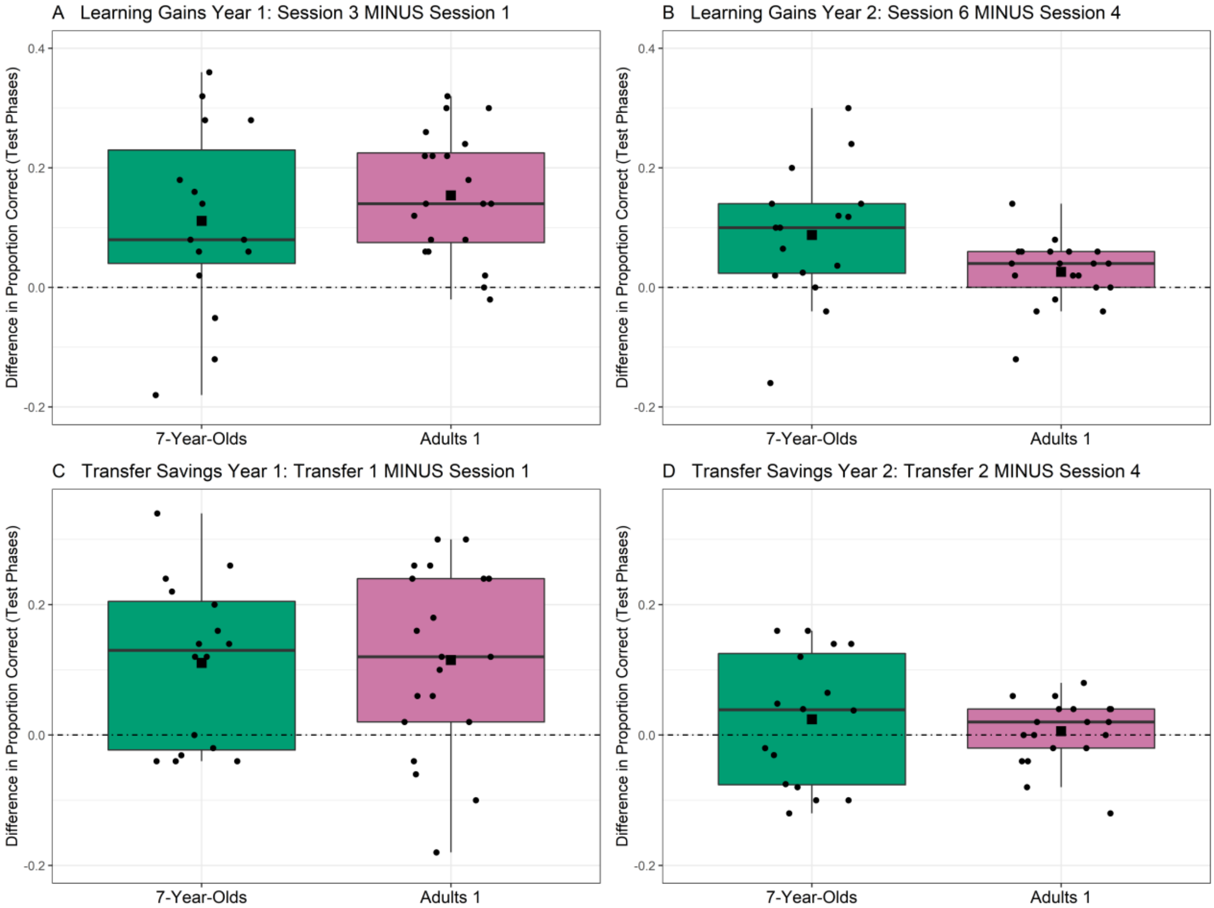
An additional analysis without the first block of each session yielded the same pattern of results (significant interactions *Year*Session* and *Age*Year* in addition to main effects of *Age, Year and Session*: all $F(1, 34) \geq 6.09$, all $p \leq .019$, all $\eta^2_g \geq .03$, all $BF_{incl} \geq 5.63$; all other $F(1, 34) \leq 3.07$, $p \geq .089$, $\eta^2_g \leq .01$, $BF_{incl} \leq .85$).

We further characterized learning across three sessions for Year 1 and Year 2 in both age groups separately by polynomial contrast analyses, which tested for linear and quadratic trends in the data (R package *emmeans*; Lenth, 2021): The increase in performance followed a linear trend in in the children and adult groups from Sessions 1 over Session 2 to Session 3 (Year 1) and from Sessions 4 over Session 5 to Session 6 (Year 2; all $p \leq .002$). In adults, an additional quadratic trend emerged over Sessions 1 to 3 (Year 1, $p = .006$).

To compare how participants transferred learned regularities to a second stimulus set in Year 1 versus Year 2 (see Fig. 7C & 7D), we conducted an ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1), *Year* (within-subject; levels: Year 1, Year 2) and *Session* (within-subject; levels: First Session [Year 1: Session 1; Year 2: Session 4], Transfer Session [Year 1: Transfer 1; Year 2: Transfer 2]).

Figure 7

Session Differences for Learning Gains and Transfer Savings for Year 1 & Year 2



Note. Learning Gains as difference in proportion correct responses of Session 3 and Session 1 (A: Year 1) or Session 6 and Session 4 (B: Year 2), respectively. Transfer Savings as difference in proportion correct responses of Transfer 1 and Session 1 (C: Year 1) or Transfer 2 and Session 4 (D: Year 2), respectively. Boxplots for 7-year-olds (green) and adults (pink) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

This analysis revealed a significant interaction of *Year*Session* ($F(1, 34) = 13.83$, $p = .001$, $\eta^2_g = .06$; $BF_{incl} = 73.70$ and significant main effects of *Age*, *Year* and *Session* (all $F(1, 34) \geq 25.30$, all $p < .001$, all $\eta^2_g \geq .09$, all $BF_{incl} > 100$; all other $F(1, 34) \leq 1.61$, $p \geq .214$, $\eta^2_g \leq .01$, $BF_{incl} \leq .52$): Overall, performance in Transfer 1 exceeded performance in

Session 1 in Year 1 ($t(35) = -5.14$, $p < .002$, $d = 0.86$, $BF_{10} > 100$; see Fig. 7C), but this was not the case for performance in Year 2 for Transfer 2 compared to Session 4 ($t(35) = -1.11$, $p = .274$, $d = 0.19$, $BF_{10} = .32$; see Fig. 7D). This effect was similar in size for both age groups (*Age*Year*Session:* $F(1, 34) = 0.18$, $p = .671$, $\eta^2_g < .01$; $BF_{incl} = .34$), Thus, no additional performance benefit for the second stimulus set was observed after relearning in Year 2.

The same pattern of results emerged from a control analysis without the first block per session (interaction *Year*Session* $F(1, 34) = 16.99$, $p < .001$, $\eta^2_g = .06$, $BF_{incl} > 100$ in addition to main effects of *Age*, *Year* and *Session:* all $F(1, 34) \geq 17.01$, all $p < .001$, all $\eta^2_g = .06$, all $BF_{incl} > 100$; all other $F(1, 34) = 1.43$, $p > .239$, $\eta^2_g = .06$, $BF_{incl} \leq .60$).

We further investigated preserved performance in the transfer session compared to the directly preceding session with the first stimulus set in Year 1 versus Year 2. To this end, we conducted an ANOVA with the factors *Age* (between-subject; levels: 7-year-olds, Adults 1), *Year* (within-subject; levels: Year 1, Year 2) and *Session* (within-subject; levels: Last Session [Year 1: Session 3; Year 2: Session 6], Transfer Session [Year 1: Transfer 1; Year 2: Transfer 2]). This analysis yielded a significant two-way interaction of *Age*Year* ($F(1, 34) = 16.86$, $p < .001$, $\eta^2_g = .03$; $BF_{incl} = 43.97$) in addition to significant main effects of *Age*, *Year* and *Session* (all $F(1, 34) \geq 8.26$, all $p \leq .007$, all $\eta^2_g \geq .03$, all $BF_{incl} \geq 14.18$; all other $F(1, 34) \leq 1.95$, $p \geq .171$, $\eta^2_g < .01$, $BF_{incl} \leq .70$): 7-year-olds displayed performance gains from Year 1 to Year 2 (Year 1 vs. Year 2: $t(31) = -4.51$, $p < .002$, $d = 0.80$, $BF_{10} > 100$), while Adults 1 did not (Year 1 vs. Year 2: $t(39) = -0.49$, $p = .623$, $d = 0.08$, $BF_{10} = .25$). For both age groups, performance declined from the last session (i.e., average of Session 3 & 6 see Table 4) to the transfer session (i.e., average of Transfer 1 & 2 see Table 4; main effect of *Session:* $F(1, 34) = 8.26$, $p = .007$, $\eta^2_g = .03$, $BF_{incl} = 14.18$). No difference in this second transfer contrast emerged when comparing Year 1 and Year 2 (*Year*Session:* $F(1, 34) = 0.30$, $p = .596$, $\eta^2_g < .01$; $BF_{incl} = .25$), independently of age (*Age*Year*Session:* $F(1, 34) = 1.95$, $p = .171$, $\eta^2_g < .01$; $BF_{incl} = .70$). Taken together, these non-significant interactions and the significant main effect of *Session* indicate that the performance of participants of both age groups decreased from the last session with the first stimulus set to the transfer session, that is the final session with a second stimulus set, in both Year 1 and Year 2.

In summary, both age groups preserved their performance level in Session 3 (last session) of Year 1 to Session 4 (first session) of Year 2 and both groups performed higher in the first session of Year 2 compared to the first session of Year 1. In Year 2, both groups improved less over three repeated learning sessions with the first stimulus set than in Year 1, and showed no additional transfer benefits for the second stimulus set from the learning sessions of Year 2. Seven-year-olds, but not Adults 1, reached higher final performance levels with the first stimulus set after relearning in Year 2, compared to Session 3 (last session) in Year 1. Adults 1 had reached ceiling in Year 1.

3.2.3. Identifying the first learning trial from modeling trial-by-trial performance

For within-session learning in Year 2, the state-space model by Smith (2005) identified the very first test trial (Block 1 in Session 1) in both age groups as the timepoint at which learning first happened. This means that at relearning, both groups showed within-session learning effects after being exposed to a single learning phase of 18 grammatical sequences. Thus, children demonstrated within-session learning effects at relearning in Year 2 as early as adults.

3.2.4. Performance correlations with explicit sequence knowledge

After *Session 6*, all 7-year-olds and all adults spontaneously mentioned to have noticed some aspect related to sequence rules when asked about the AGL task and their decision strategies about stimulus set 1 (open questions, see *Explicit Knowledge of Sequence Rules;* all answers see Table 3). After *Transfer 2,* the same pattern occurred with all participants reporting to have become aware of sequence rules.

Scores in reported explicit sequence knowledge in Year 2 (assessed at Transfer 2 about specific legal bigram transitions, see specific questions in *Explicit Knowledge of Sequence Rules*) did not significantly differ between 7-year-olds and Adults 1 ($t(34) = 0.43$, $p = .249$, $d = 0.18$, $BF_{10} = .35$).

7-year-olds reported more explicit knowledge about sequence rules at the end of Year 2, compared to the end of Year 1 ($t(15) = -3.51$, $p = .004$, $d = 0.90$, $BF_{-0} = 56.33$; one-sided), while this difference between Year 1 and 2 did not reach statistical significance in Adults 1 ($t(19) = -1.43$, $p = .084$, $d = 0.32$, $BF_{-0} = 1.02$; one-sided). For all associations of explicit sequence knowledge with AGL Learning Gains and Transfer Effects (7-year-olds: all $| r_s | \leq .31$; $p \geq .472$, $BF_{10} \leq .68$; Adults 1: all $| r_s | \leq .38$, $p \geq .180$, $BF_{10} \leq 1.07$), see Appendix B.

Task difficulty and short-term familiarity caused similar performance effects in Year 2 as in Year 1 (see Appendix B).

4. Discussion

The goal of Project 1 was to characterize long-term learning trajectories in a modified visual AGL task in children and adults across one year. It was tested whether 7-year-olds outperform adults in artificial grammar learning and whether this advantage translates to superior long-term memory of AG rules. To this end, both 7-year-olds and adults engaged in a multisession visual AGL task which was repeated one year later.

We found successful AG learning and transfer of rule knowledge to another stimulus set in both 7-year-olds and adults. However, adults learned quicker and overall performed at a higher level. Both groups retained rule knowledge over a period of one year, started at the level reached one year earlier and continued to improve, although to a lesser degree than in the first year. We did not find better retention effects or relearning advantages across the second set of sessions after the one-year delay in 7-year-olds compared to adults. Explicit knowledge of the AG rules was indistinguishable between adults and children and correlated with transfer gains in children.

In the following, discussions will focus on a child-adult comparison in sequence learning across two timescales (1 week & 1 year). Implications for more general concepts like sensitive periods in development will be discussed in Chapter V.

4.1. 7-year-olds acquire AG rules but overall learn slower and perform worse than adults across several sessions

When exposed to a visual AGL task, 7-year-olds and adults learned the visual sequence rules. Adults overall outperformed children, but learning gains over several sessions were indistinguishable between both groups. These results are in accord with Ferman and Karni (2010) and Smalle, Page, et al. (2017), who reported indistinguishable learning rates across multiple sessions in 8-12-year-olds and adults in phonological sequence learning tasks. Our study extends these findings to visual sequence learning with more complex rules as defined in an AG.

Our study allowed for a direct comparison between both age groups with regard to short-term retention effects, which elaborates evidence on visuomotor retention across 24 hours in an alternating serial reaction time task for 9-15-year-old children (Tóth-Fáber, Janacsek, & Németh, 2021) and adults (Kóbor et al., 2017) from separate studies. Our finding that 7-year-olds and adults continued learning at their last performance level after delays of at least one night between the sessions of Year 1 suggests similarly effective short-term consolidation processes for learned regularities in children and adults. We can only speculate on the role of sleep for performance in our task setting. Sleep-dependent consolidation of encountered environmental patterns has been proposed to rely mainly on time-compressed replay processes in the hippocampus during slow-wave sleep (Lerner & Gluck, 2019; Wilhelm et al., 2012), which enable the long-term storage of rule knowledge in cortical networks and consequently the generalization of the acquired rules.

In contrast to previous findings on rule generalization in phonological single-session (Hickey et al., 2019) and multi-session (Ferman & Karni, 2010) learning, we found that children as young as seven years of age generalized their visual rule knowledge to the same extent as young adults did. Transfer effects in our study emerged without any explicit instructions, i.e. without providing participants explanations about the nature of these rules. This finding, thus, is incompatible with the proposition that explicit instructions are necessary for successful transfer in children younger than 12 years of age (Ferman & Karni, 2010, 2014). In line with our findings, in a single-session study, 6- to 9-year-olds were able to implicitly extract categorical regularities in visual triplet learning and demonstrated a transfer of rule knowledge to unseen items of the same picture category to the same degree as adults (Jung et al., 2020). Even younger children with age 3 to 6 years were reported to successfully transfer learned regularities to new syllables in auditory artificial grammar learning, when underlying rules reflected word form distributions from natural language (Nowak & Baggio, 2017). These inconsistent findings raise the question of under which conditions children transfer learned regularities to new items or even to new categories. In line with previous suggestions (Gomez, 1997; Nowak & Baggio, 2017; Witt et al., 2013), we hypothesize that in learning complex regularities, task features that help building explicit knowledge about these regularities might favor rule abstraction and consequently rule transfer to new surface features of the instantiated sequence rules in children. In the present study, task features that promoted the acquisition of explicit knowledge possibly included instructing participants in a way that linked learning and test phases (Witt et al., 2013) and providing audio-visual feedback after each test trial (Nowak & Baggio, 2017). This link between explicit rule knowledge and rule transfer is supported by our finding that in children, larger Transfer Savings (Transfer 1 minus Session 1) showed a trend to be positively associated with more explicit knowledge of the rules at the end of Year 1.

Despite the similarities in learning trajectories and transfer effects for children and adults, we found age differences within the first year of learning across one week: Adults needed less exposure to the AG within one session to successfully apply the sequence rules in subsequent test trials: After only one learning phase with 18 grammatical sequences they performed above-chance performance. By contrast, successful learning in children was not observed before the third test phase of Session 1, i.e., only after exposure to 54 grammatical sequences. Adult’s learning curves furthermore showed an exponential increase across three sessions with the same stimulus material. In contrast, children’s learning curves increased linearly. Interestingly, these age differences were no longer observed for relearning in Year 2. Instead, children in Year 2 showed within-session learning as early as adults (successful learning was observed after a single learning phase) and a linear increase of performance emerged in both age groups. The statistical learning model of Janacsek et al. (2012) (further elaborated by Daltrozzo & Conway, 2014; Janacsek et al., 2012; Nemeth et al., 2013) has suggested a switch to a more supervised, “model-based” learning system, taking place in the course of development. Model-based learning was considered to rely more on attentional resources, behavioral control and prior knowledge than the “model-free” learning, which has been proposed to be the predominant form of learning in early development. The reliance on supervised learning mechanisms in sequence learning might allow adults to reach a higher performance levels at test after exposure to sequential regularities (discussed as explicit learning markers, see Chapter I & Chapter V, based on Forest et al., 2023). By contrast, event-related potential (Jost et al., 2011) and reaction time (Janacsek et al., 2012) studies have suggested that children better implicitly pick up regularities and seem to do so at an earlier timepoint during acquisition than adults (discussed as implicit learning markers, see Chapter I & Chapter V based on Forest et al., 2023).

Additionally, adults and children in the present study did not significantly differ in their reported levels of explicit sequence knowledge by the end of the transfer session. Previous studies have mainly relied on verbal reports about legal transitions or sequence rules (Ferman & Karni, 2010; Hickey et al., 2019; Smalle, Page, et al., 2017), production tasks of legal sequences (Jung et al., 2020) or confidence ratings at test (Smalle, Page, et al., 2017) to measure explicit knowledge and found that adults had acquired more knowledge about sequence rules than children. The lack of a significant association between explicit knowledge and sequence learning in adults might be due to ceiling performance in this group. In fact, ceiling performance was observed as early as Session 2. The lack of age differences in explicit knowledge might be due to the fact that we assessed explicit sequence knowledge only after four sequence learning sessions on separate days had been completed, in order to avoid inducing a change in learning strategies. Wilhelm et al. (2013) tested the explicit recall of sequential transitions from learning a deterministic sequence in an implicit visuomotor task after a sleep vs. wake phase. They found that children aged 8-11 years benefitted more than adults from sleep in acquiring explicit knowledge from their implicit learning experience.

4.2. Children and adults retain visual regularities over a one-year delay and show comparable relearning effects after the delay

From Year 1 to Year 2, both age groups retained their final level of rule knowledge over the 12-month break, confirming the hypothesized long-term retention of sequential regularities in adults and children. This corroborates findings on long-term retention of sequence knowledge across two months (Ferman & Karni, 2010) and one year (Smalle, Page, et al., 2017) in the phonological domain. These studies enclosed both 8- to 12-year-old children and adults. Our results on retention of complex visual regularities furthermore extend findings from visuomotor retention across a one year delay on 9-15-year-olds' (Tóth-Fáber, Janacsek, & Németh, 2021) and adults' learning (Kóbor et al., 2017) from separate studies: We were able to directly compare children and adults and show that adult-like retention is already evident in children as young as 7 years. When comparing children's and adults' learning within the same task setting, Ferman and Karni (2010) reported similar long-term retention effects in accuracy for 8-12-year-old children and adults. Smalle, Page, et al. (2017) found no age differences in retaining an announced and cued syllable sequence ("explicit sequence"), but an advantage of 8-9-year-olds vs. adults in retaining an implicitly learned syllable sequence ("implicit sequence") over a 12-month break when they matched baseline performance levels before the 12-month break between both age groups. In the same vein, we found a trend for better one-year retention in children (performance increment from Session 3 to 4) compared to adults (performance decrement from Session 3 to 4), when analyzing performance without the first task block of each session. Leaving out the very beginning of these sessions probably mitigated the adult advantage over children to draw on their general resources of greater attention and cognitive control, which should be most advantageous in a rather new or less practiced task setting (Session 4 after 1-year delay). This demonstrates children's ability to retain visual regularities over one year to at least the same degree as adults, despite their overall lower memory capacities (Gathercole, 1998).

In addition to testing retention of complex visual regularities after a long-term delay in a single follow-up session, we extended the existing literature by investigating how both age groups use their acquired rule knowledge in another set of multiple relearning sessions in Year 2: Children and adults continued to improve with the first stimulus set, although to a lesser degree than in Year 1. This was true for both age groups, rendering the explanation unlikely that this decrease in learning gains was solely due to ceiling effects preventing further performance improvements in Year 2: Children were performing at an average of 71% correct in the first session of the second year (Session 4), leaving enough room for them to further improve in the sessions to come. Nevertheless, they did not improve as much over the consecutive three sessions (Session 4 to 6) as they did in Year 1 (Session 1 to 3). Adult performance was more constrained, because they had already reached ceiling in Year 1.

We cannot exclude the possibility that in Year 2, children's performance was influenced by better sequence learning abilities compared to Year 1, due to trivial maturational effects (e.g., of memory capacity).Previous studies (Arciuli & Simpson, 2011; Raviv & Arnon, 2017; Shufaniya & Arnon, 2018) have demonstrated that children's behavioral performance in visual triplet learning tasks improves between 5 and 12 years of age. Nevertheless, it is less clear how an age difference of only one year (in our sample, children were 7 years old in Year 1 vs. 8 years old in Year 2) impacts the acquisition of sequential regularities. Available data for 6.5-7.5-year-olds vs. 8-9-year-olds from the cross-sectional study by Raviv and Arnon (2017) show inconsistent trends for different stimulus material, with younger children performing better in auditory/linguistic triplet learning vs. older children performing better in visual/non-linguistic triplet learning. We think that two observations in our data speak against the proposition that relearning effects in children reported here can be sufficiently explained by maturational effects: First, children preserved their last performance level across one year (Session 3 to Session 4), corroborating the phenomenon of experience-dependent benefits for relearning in school-aged children reported in previous studies (Smalle, Page, et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021). Second, the relearning gain with the same stimulus material over three sessions in Year 2 (Session 4 to 6) compared to the learning rate in Year 1 (Session 1 to 3) was lower. If children were better at acquiring regularities due to their more matured cognitive abilities in Year 2, they would have been expected to show a greater increase in learning in Year 2 compared to Year 1, however. Importantly, all relearning effects were confirmed after eliminating the first task block of Session 1 of Year 1 and Session 4 (first session) of Year 2, respectively, to exclude trivial task familiarity effects.

Both age groups did not show additional transfer to the second stimulus set in Year 2. This suggests that the acquired rule knowledge from Year 1 was applied to the same extent in the more frequently encountered (stimulus set 1) and the less frequently encountered (stimulus set 2) learning material, benefitting performance in both learning situations alike. Since additional experience with the first stimulus material further improved performance in Year 2 (Session 4 to 6) in both age groups, we think it is unlikely that additional transfer gains in Year 2 (Session 4 to Transfer 2) were prevented by ceiling effects. It has been argued that offline periods promote the extraction and representation of underlying regularities by replay-induced strengthening of memory representations (Lerner & Gluck, 2019; Y. Liu et al., 2019; Wilhelm et al., 2012), thereby enabling transfer effects in learning at the level of cortical circuits. There is little evidence on how these neural mechanisms operate over extended time periods like the 12-month period employed in the present study. Non-human animal studies (Xu et al., 2009; Yang et al., 2009) have reported persistent structural changes in the cortex (i.e., proliferation and reactivation of dendritic spines), when rodents acquired new sensory or motor skills and relearned them after a long delay. This could provide a possible explanation how the neural architecture implements long-term learning effects, as the previously acquired task-specific neural infrastructure might be used more efficiently when reencountering the same learning environment (discussed in Hofer & Bonhoeffer, 2010). This mechanism of structural plasticity can be speculated to underlie relearning effects reported in our study and in previous behavioral reports on human relearning (Livosky & Sugar, 1992; Murre & Dros, 2015; Parkin & Streete, 1988), which date back to Ebbinghaus' observations on faster and more efficient learning of familiar as opposed to new material (Ebbinghaus, 1880).

4.3. Conclusion

Project 1 showed successful AG learning and transfer of visual rule knowledge to another stimulus set in both 7-year-olds and adults, but adults learned quicker and overall performed at a higher level. We report that both age groups successfully use their retained rule knowledge after one year, which extends studies on retention by characterizing relearning across another set of multiple relearning sessions after a delay. However, we did not observe relearning advantages across the second set of sessions after the one-year delay in 7-year-olds compared to adults. Explicit knowledge of the AG rules was indistinguishable between adults and children and correlated with transfer gains in children. These findings challenge the notion that a more implicit extraction of sequential regularities early in life results in better learning outcomes in the long run. Project 2 will look into learning trajectories of children younger than 7 years, and addresses the question how age-related maturation vs. genuine effects of prior learning influence relearning after a delay.

Chapter III: Repeated visual statistical learning - Do younger children show long-term learning advantages (Project 2)?

1. Introduction

Extracting sequential regularities from the environment has been suggested to underly, for example, language acquisition (Conway & Pisoni, 2008; Deocampo et al., 2018; Erickson & Thiessen, 2015; Romberg & Saffran, 2010) and motor skill learning (Lukács & Kemény, 2015; Savion-Lemieux et al., 2009). This mechanism has been investigated as statistical learning or implicit learning and is referred to as sequence learning throughout this dissertation (see Chapter I). A controversy has emerged, however, as to whether sequence learning abilities improve across development (Arciuli & Simpson, 2011; Raviv & Arnon, 2017; Schlichting et al., 2017; Shufaniya & Arnon, 2018), or are best early in live and then decrease from middle childhood to adulthood (Janacsek et al., 2012; Jost et al., 2011; Nemeth et al., 2013; Rohlf et al., 2017; see Chapter I for a comprehensive review). For telling apart rule-following from rule-violating sequences in the visual domain, discrimination performance has been shown to improve in the age range from 5 to 12 years (Arciuli & Simpson, 2011; Raviv & Arnon, 2017; Shufaniya & Arnon, 2018), which was extended to the age range from 6 to 30 years by Schlichting et al. (2017) (see also Lukács & Kemény, 2015; Weiermann & Meier, 2012 for improved performance with age in other skill-based and probability learning tasks). At the same time, there are reports of a high initial sensitivity for environmental regularities early in life, which decreases later in childhood: This is mainly reflected in neural markers (like ERPs) during passive presentation, which, e.g., provided evidence for extracting cross-modal, audio-visual statistics from mere exposure in infants, but not in adults (Rohlf et al., 2017). An early sensitivity towards environmental regularities has additionally been shown in childhood, using “online” behavioral learning markers (e.g., reaction times for mapping motor responses to a visual sequence of stimulus locations vs. random locations): Reaction time improvements for visual sequences were shown to be the greatest in the youngest age groups of 4 to 12 year-olds (Janacsek et al., 2012) in a large cross-sectional sample up to 85 years of age.

Others have tried to accommodate these two, seemingly contradictory, age trends in sequence learning by sorting available evidence along two main, interrelated dimensions (for a more thorough taxonomy, see Chapter I): Firstly, Forest et al. (2023) proposed that reported age changes mainly differ depending on whether direct vs. indirect learning markers of sequence learning were measured. When focusing on indirect measures like ERPs or reaction time improvements (both assessed without any “overt report” from participants, as defined for indirect markers by Forest et al., 2023, p. 3), younger age groups tend to show stronger than or at least comparable learning effects to their older counterparts (Janacsek et al., 2012; Jost et al., 2011; Nemeth et al., 2013; Rohlf et al., 2017). In contrast, older children and adults seem to outperform younger children in more direct measures of learning, which are mainly based on behavioral performance in a test phase after exposure to sequential regularities (Arciuli & Simpson, 2011; Raviv & Arnon, 2017; Schlichting et al., 2017; Shufaniya & Arnon, 2018; see Forest et al., 2023, however, for the special case of linguistic stimulus material). Secondly, contributions of distinct neurocognitive systems for implicit vs. explicit learning mechanisms have been proposed to shift towards explicit learning across development (Conway, 2020; Daltrozzo & Conway, 2014; Nemeth et al., 2013). This “natural” shift might contribute to the observed decrease in tracking (passively) encountered sequential regularities across infancy and childhood (based on declining reliance on the implicit system), while at the same time growing better at deciding what makes up a legal vs. an illegal sequence (driven by increasing reliance on the explicit system).

Apart from this developmental shift, Batterink et al. (2015) argue that the explicit system can be deployed as an additional resource for, e.g., more complex input, complementing implicit mechanisms that constitute the “default mode” in sequence learning. This might add to variations in learning outcomes as a function of the situation at hand, as soon as explicit components (e.g., greater selective attention) are developmentally available to a learner. Relatedly, H. Liu et al. (2023) argue that implicit and explicit memory traces emerge simultaneously (see also Batterink et al., 2015; Conway, 2020), but can be dissociated by using indirect (implicit) vs. direct (explicit) learning markers. In their study, they tested how fast implicit vs. explicit memory traces of sequential regularities decay and if they are influenced by repeated testing. While an adult sample acquired and reactivated these two types of representations in parallel, respective representations seemed to be consolidated differently across 24 hours: Explicit memory traces decayed faster, but became more abstract with time and were more strongly influenced by later testing than implicit memory traces (H. Liu et al., 2023).

So, by distinguishing direct vs. indirect learning markers (Batterink et al., 2015; Conway, 2020; Forest et al., 2023), while at the same time mapping contributions of implicit vs. explicit neurocognitive mechanisms (Conway, 2020; Daltrozzo & Conway, 2014; Nemeth et al., 2013), findings of both, better vs. declining, sequence learning abilities with age can be accommodated. However, this framework is mainly based on age differences measured in single sessions of sequence learning, limiting the ecological validity of the modeled learning processes across time. This means, it remains unclear how exactly the proposed developmental make-up influences sequence learning outcomes in the long run, when the acquisition and use of sequential regularities are assessed over several learning instances and after a long-term delay.

Multi-session studies so far have mainly considered children aged 8 years and older in comparison to adults (Ferman & Karni, 2010, 2014; Smalle, Page, et al., 2017), investigating how they improve in learning spoken syllable sequences across several task encounters. In these studies, auditory learning rates have been shown to increase across sessions to a comparable degree in children age 8 to 12 years and in adults, with adults overall outperforming children (Ferman & Karni, 2010, 2014; Smalle, Page, et al., 2017) and 12-year-olds overall outperforming 8-year-olds (Ferman & Karni, 2010). It is less clear how children under 8 years of age differ in their multi-session acquisition rates of sequential regularities. The only available data for younger children with age 6 and 7 years comes from studies on visuo-motor sequence learning with a single follow-up session after 24 hours (Juhász & Németh, 2018; Savion-Lemieux et al., 2009). Savion-Lemieux et al. (2009) investigated sequence learning of three child groups aged 6, 8 and 10 years in addition to adults: Performance accuracy for visuo-motor associations of a deterministic sequence improved to the greatest extent in younger children (6-year-olds = 8-year-olds > 10-year-olds = adults), measured as correctly finger-tapping repeating screen-locations in the end of Day 2 relative to the beginning of Day 1 (see Fig. 5 in Savion-Lemieux et al., 2009). However, in their comprehensive cross-sectional study, Tóth-Fáber et al. (2023) utilized a more complex visuo-motor sequence learning task in order to prevent ceiling effects and report age-independent retention effects in 9 age groups from 7 to 76 years across a 24-hour delay. In accord with this, Juhász and Németh (2018) reported age-related changes in within-session acquisition rates for visuo-motor sequence learning in the same task, but no age differences in retention across a 24-hour-interval for six age groups ranging from 7 to 29 years. So, for developmental samples below the age of 8 years, the use of sequential regularities across multiple sessions separated by short-term delays has yet to be characterized. It remains to be tested if adult-like performance improvements and retention rates, documented in previous studies, hold for younger children and might even extend to longer delays.

Concerning retention and consolidation rates of sequential regularities across longer delays of several months, previous research has focused on children older than 8 years as well. Evidence for this age group suggests that eliciting mainly implicit vs. rather explicit processes might matter for observing age differences in retention. No age differences were reported when task protocols included some explicit component like cueing sequences, explicitly stated sequence rules prior to learning, or provided performance/visuomotor feedback: Implementing respective task features, children (age 8-12 years, Ferman & Karni, 2014; Smalle, Page, et al., 2017) and adults retained auditory (syllable) sequences equally well across delays of two months and one year, respectively. In the same vein, sequence knowledge one year after the last learning session was not associated with age within the investigated group of children age 9 to 15 years in visuo-motor learning (Tóth-Fáber, Janacsek, & Németh, 2021). By contrast, for long-term retention of implicit sequence knowledge (i.e., acquired without any of the described task features from above), one study has reported better retention up to 12 months in children of age 8-9 years vs. adults (Smalle, Page, et al., 2017).

Including children younger than 8 years when mapping long-term trajectories of sequence learning seems to be of particular interest, considering language and memory development in this age range: The literature on sensitive periods in language learning proposes that early language learning critically shapes later learning (Werker & Hensch, 2015), with most efficient grammar learning likely to take place before the age of 7 (J. S. Johnson & Newport, 1989). The concept of sensitive periods, enabling better learning at a young age, has recently been translated to timelines of childhood advantages in other cognitive domains. These include learning probabilistic information and recalling object locations, properties and associations (Gualtieri & Finn, 2022). Several of these documented advantages of (younger) children in learning and remembering center around the age of 4 to 7 years, a developmental time that has been identified to entail great changes in general learning mechanisms (Sameroff & Haith, 1996), related e.g., to cognitive control (M. H. Johnson & Munakata, 2005; Ramscar & Gitcho, 2007). The continued development of cognitive control and prior knowledge factors into how information from several learning experiences are integrated into memory (Brod et al., 2013). Children were reported to integrate these experiences late in the memory process by making inferences at retrieval, while adults may perform this integration as early as encoding (Shing et al., 2019). Thus, the above literature warrants an investigation of children of different ages in multi-session sequence learning, preferably younger than 8 years, in addition to child-adult comparisons.

Adopting multi-session paradigms that include a long-term delay allows to test age-related changes in the step-by-step acquisition and subsequent use of sequential rule knowledge.

In addition, it is vital to ask if and how development influences the degree to which learning experiences can be generalized to new input. Memory processes seem to shift from emphasizing generalization in early childhood to stronger memory specificity until middle childhood (Keresztes et al., 2018; Ngo et al., 2018). This literature implies that memory processes change towards more specific encoding and retrieval of object representations and object-object associations around the age of 6 years. Relatedly, memory representations from sequence learning have been proposed to become increasingly specific until age 6-7 years (Forest et al., 2021; Forest et al., 2023), corresponding roughly to the same age range. More pronounced forgetting between several learning experiences early in development could additionally favor higher transfer abilities in younger children, as suggested by the forgetting-by-abstraction account (Vlach, 2014; elaborated in Chapter I). The role of forgetting in sequence learning has recently been stressed by Endress and Johnson (2021) in computational modeling to further the mechanistic understanding of the involved processes, and by Forest et al. (2023) for developmental predictions (see Chapter I). Behavioral evidence that directly compares transfer effects in sequence learning between children of different ages (3-12 years) and adults remains inconclusive, however (Ferman & Karni, 2010, 2014; Jung et al., 2020; Nowak & Baggio, 2017). In the only multi-session study available, children of age 8 years (as opposed to children aged 12 years and adults) failed to generalize a learned sequence rule to new stimuli (Ferman & Karni, 2010), except when being told what constitutes the sequence rule before exposure (Ferman & Karni, 2014). Notably, these findings on transfer effects in sequence learning across development are not only diverging, but fall short of speaking to learning in children younger than 8 years in task protocols with several sessions and across longer delays between sessions (> 2 months). Children around age 6 provide a promising sample for testing transfer effects in these settings, given the proposed shift in memory processes and representations around this age as elaborated above.

Our study is the first to look into multi-session learning of visual sequences in two child groups of age 5 and 6 years across 2 sets of sessions with a long-term delay of 12 months in between. The aim of the study was to evaluate how development influences the repeated use of acquired sequence knowledge. The child groups of age 5 years and 6 years were chosen because, as discussed in the previous paragraphs, the age range of 4 to 7 has been identified as a period of change in learning mechanisms from the language, memory and generalization literature. In addition to testing the effects of different onsets of multiple learning experiences in childhood (at age 5 vs. at age 6 years), learning outcomes of both child groups were related to adult performance. Furthermore, we included a generalization test of the learned sequence rules to new visual surface features. This allowed us to explore how abstract the acquired sequence knowledge is represented at different ages, i.e., if any age differences emerge when this knowledge is accessed and used for new learning material that displays the same underlying rules.

Project 2 implemented the same AGL task with visual stimuli and complex sequence rules as Project 1, explained in detail in Chapter II. This sequence learning task was completed by three age groups, 5-year-old children, 6-year-old children and adults (Adults 2, see Table 1 and Figure 8). All participants learned in three sessions over the course of one week; After one year, first remaining sequence knowledge was tested in three "relearning" sessions with the original item set and subsequently transfer to a new visual stimulus set was tested in the final session. To disentangle maturational effects at relearning after one year from the effects of prior learning, we included a comparison to naïve groups of children of the same age as 5-year-olds and 6-year-olds after the delay (i.e., 6- & 7-year-old controls), for whom the AGL task was completely unfamiliar (see Table 1).

We hypothesized to find an age-independent increase in sequence learning performance across sessions in both child groups and Adults 2 for the first stimulus set, over the course of one week (Year 1). After the one-year delay, we expected to observe preserved AG knowledge as well as transfer effects at the end of the second set of sessions (Year 2). Based on an early childhood advantage reported as a higher sensitivity towards sequential regularities, greater forgetting that should promote the extraction of abstract regularities, and stronger overgeneralization in the domains of memory and language, younger age groups were expected to feature higher retention of the acquired rule set over one year, to quicker implicitly relearn the AG, and to show larger transfer to a new stimulus set (5-year-olds > 6-year-olds > Adults 2).

Additionally, we expected that 5- and 6-year-olds predominantly rely on implicit knowledge, while adults acquire more explicit knowledge about the underlying sequence rules (Ferman & Karni, 2010; Hickey et al., 2019; Jung et al., 2020; Smalle, Page, et al., 2017). If any trend across the restricted age range we covered could be expected, the older children were thought to acquire more explicit knowledge (5-year-olds < 6-year-olds) and possibly depend more on their acquired knowledge for improvement and transfer in sequence learning.

2. Methods

2.1. Participants

Project 2 involved two groups of healthy children, 35 five-year-olds (5 years old ± 2 months at Session 1) and 34 six-year-olds (6 years old ± 2 months at Session 1) from the City of Hamburg, Germany. Additionally, an adult group was recruited which consisted of 32 healthy adults, mostly undergraduate students recruited from the University of Hamburg. None of the participants reported a history of seeing or hearing impairments, nor any neurological disease. They all were native German speakers. The group of 7-year-olds was additionally included as a control group for 6-year-olds in Year 2 and is described in detail in Chapter II.

To maximize on all available data in this multi-session study design, the different types of analyses – on session data for both years vs. only for one year vs. on trial-wise data per subject – included slightly different numbers of participants from these age groups as described below and detailed in the respective results sections. For an overview of the included datasets for each analysis, sample sizes per age group are listed in the tables detailing the subsequent analyses in the beginning of each results section.

For all analyses requiring data from Year 1 and Year 2 (see results section *Performance in Year 2 benefits from prior learning in Year 1*), the data of 11 five-year-olds, 7 six-year-olds and 12 adults had to be excluded from the analyses, due to missing data. Eighty-three participants returned for the second set of sessions in Year 2 (drop-out of 3 five-year-olds, 4 six-year-olds and 11 adults), but we additionally had to exclude 8 five-year-olds and 3 six-year-olds of the returning participants since they had missing data in more than half a session across both years (≥ 3 task blocks in at least one session of both years). One additional adult had to be excluded due to technical failure leading to data loss. The remaining 24 five-year-olds (15 female, mean age at Session 1: 5.08 ± 0.09 years, range: 4.91-5.22 years), 27 six-year-olds (11 female, mean age at Session 1: 6.05 ± 0.09 years, range: 5.83 – 6.16 years), and 20 adults (17 female, mean age at Session 1: 24.73 ± 6.45 years, range: 19.21-49.70 years) were included for all analyses requiring data from Year 1 and Year 2 (see results section *Performance in Year 2 benefits from prior learning in Year 1*). All participant characteristics for this final sample are listed in Table 5. Of the additional group of 7-year-olds from Project 1, we also included all children with available data for all sessions in Year 1 and Year 2 (16 seven-year-olds, described in Table 3 of Chapter II) to provide a comprehensive overview over all child groups who were part of this dissertation, as explained below (see results section *Improvement and transfer in all child groups including 7-year-olds*).

For relearning analyses on data of Year 2 in the child groups (see results section *Relearning advantages in Year 2 for 5-year-olds and 6-year-olds compared to naïve controls*), we included all available datasets for the age-matched naïve control groups, who fulfilled the missing data criterion from above (≥ 3 task blocks in at least one session of Year 1 or 2, respectively): This amounted to data of 31 six-year-olds from Year 1 to match 24 five-year-olds in Year 2 and data of 27 seven-year-olds from Year 1 to match 27 six-year-olds in Year 2. The group of 7-year-olds is described in detail in Chapter II (Table 2).

For analyses requiring single trial responses of Session 1 to 3 in Year 1 and Session 4 to 6 in Year 2 (see results sections *Earlier trial-by-trial learning effects in Year 2 compared to Year 1*, *Earlier trial-by-trial learning effects in Year 2 due to prior learning in 5-year-olds and 6-year-olds compared to controls*), we included all complete datasets of 24 five-year-olds, 25 six-year-olds and 20 adults. The additional age group of 7-year-olds included n = 27 complete datasets for Session 1 to 3 in Year 1 and n = 15 complete datasets for Session 4 to 6 in Year 2 (described in section *Participants* in Chapter II).

All procedures for participant compensation and ethics approval were described in the *Methods* section of Chapter II.

Table 5

Participant Characteristics for 5-Year-Olds, 6-Year-Olds, Adults 2 (Included Datasets)

Participant Characteristics	**5-Year-Olds (*n* = 24)**	**6-Year-Olds (*n* = 27)**	**Adults 2 (*n* = 20)**
Time period betw. Year 1 & 2 (months betw. Session 1 & 4)	12.38 (1.21) [11.00-15.00]	11.85 (1.03) [12.00-13.00]	11.45 (0.51) [11.00-12.00]
Days betw. Sessions of Year 1			
Session 1 to 3	5.46 (1.32)	5.04 (1.76)	3.45 (1.36)
Days betw. Sessions of Year 2			
Session 4 to 6	4.29 (1.37)	5.04 (2.31)	4.20 (1.54)
Session 4 to Transfer 2	7.17 (1.34)	7.15 (1.90)	6.85 (1.79)
Age Year 1 (years)	5.08 (0.09) [4.91-5.22]	6.05 (0.09) [5.83 – 6.16]	24.73 (6.45) [19.21-49.70]
Age Year 2 (years)	6.15 (0.13) [5.93-6.42]	7.06 (0.07) [6.97 – 7.18]	25.97 (7.88) [20.20-50.68]
Gender (f/m)	15/9	11/16	17/3
School/Education [a]	*n* = 20 kind *n* = 4 pre	n = 6 kind *n* = 20 pre *n* = 1 scho	*n* = 19 university students
Bilinguals	3	4	4
Daily mobile device usage [a] (min)	12.41 (11.90)	29.23 (34.48)	381.43 (167.64)

Note. M (SD) [range]; betw. = between, kind = kindergarten, pre = preschool, sch = school, univ = university.

[a] assessed in Year 1 (Session 1).

2.2. Design and procedure

2.2.1. Study design

All participants completed a total of three sessions in Year 1 on separate days (Session 1, 2 & 3), and an equivalent of three sessions (Session 4, 5, 6) with a subsequent transfer session (Transfer 2) in Year 2, each set spread out over the time of approx. one week (see study design Fig. 8 and Table 5 for information on session timing of all age groups), in which they completed a visual sequence learning task.

- Session 1 & 4 (in the lab): After the assessment of working memory, the first learning session with the tablet computer (with stimulus set 1) followed. Next, we measured declarative memory and German grammar skills (see *Memory and language skills*). Session 1 & 4 lasted 90 to 120 minutes each including participant briefing and breaks.
- Session 2-3 & 5-6 (at home): Two more learning sessions took place with stimulus set 1 on the tablet computer.
- Transfer 2 (in the lab): The second (new) stimulus set was introduced on the tablet computer, but with the same underlying rule set to test transfer of AG learning. Moreover, explicit knowledge about sequence orders was assessed with a questionnaire in adults and adapted questions with picture cards in children (see *Explicit knowledge of sequence rules*).

Due to health regulations related to the COVID-19 pandemic, some participants completed more than the planned two sessions in Year 2 at home (n = 3 five-year-olds, n = 2 six-year-olds and n = 2 Adults 2, all with 3 or all 4 sessions as at home sessions). For these sessions, participants received all material and a tablet computer by mail and completed the respective learning sessions at home.

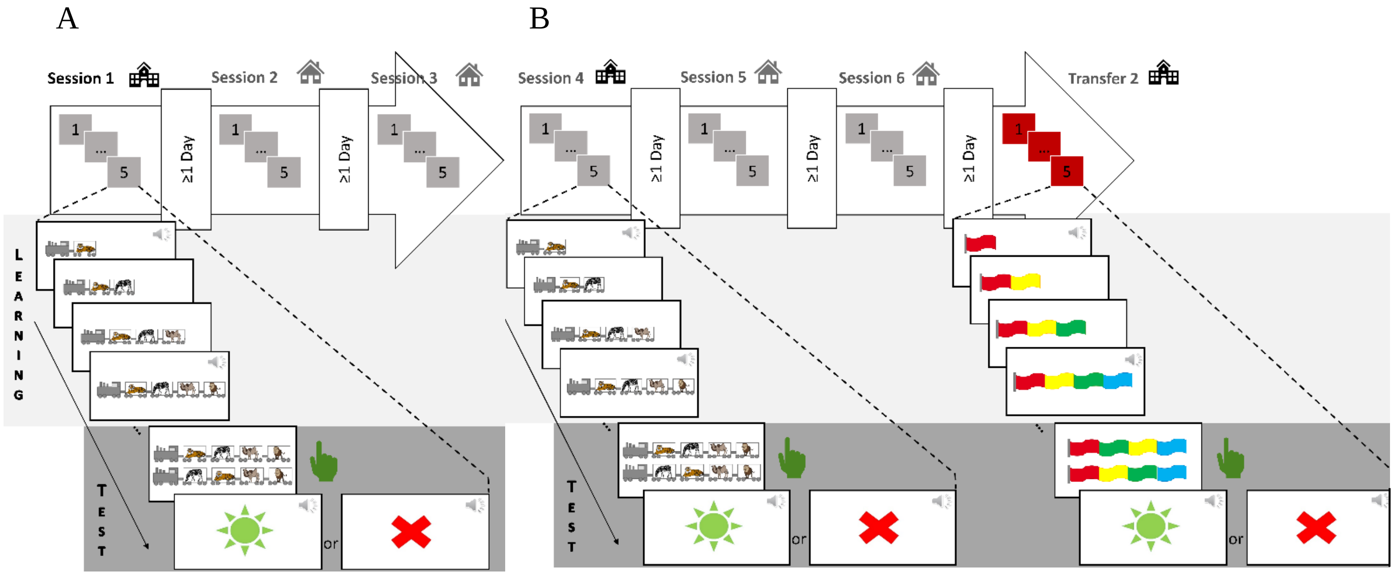
2.2.2. Visual sequence learning task

Details of the modified AGL task to measure multi-session visual sequence learning are provided in Chapter II (*Stimuli and apparatus*). It implemented the artificial grammar system introduced by Reber (1967) for constructing grammatical sequences (see Fig. 2). The task entailed 5 task blocks per session, consisting of alternating phases of learning (exposure to 18 grammatical sequences each) and test (10 trials of two-alternative forced choice responses between one grammatical and one ungrammatical sequence). This amounted to an AGL task exposure of approx. 25-30 minutes per session. The first 3 sessions of each year used a first stimulus set (stimulus set 1), while the last session Year 2 employed a second stimulus set (stimulus set 2) to investigate transfer of learned AG rules.

As shown in Project 1, children of age 7 years and adults did not differ systematically in their performance on the two dimensions of test trials, difficulty and short-term familiarity (explained in *Construction of grammatical and ungrammatical sequences* in Chapter II). Thus, in all analyses of the current Project 2, responses were collapsed across these dimensions to reduce complexity for the extensive body of analyses reported here. However, equivalent analyses to Project 1 are reported in Appendix C for the age groups investigated here, which tested the overall influence of both trial dimensions of test trials (i.e., "easy" vs. "difficult" trials, "seen" vs. "not seen" trials, respectively) on group performances of 5-year-olds, 6-year-olds and Adults 2, separately for Year 1 and Year 2.

Figure 8

Study Design of all 7 Sessions of Visual Sequence Learning in Year 1(A) & Year 2 (B)



Note. The first 3 sessions of each year used a first stimulus set (here: Stimulus Set Animals), while the last session of Year 2 employed a second stimulus set (here: Stimulus Set Colors) to investigate transfer of learned AG rules. Each session consisted of 5 task blocks with alternating learning (light gray background) and test (dark gray background) phases.

2.3. Material

2.3.1. Explicit knowledge of sequence rules

Explicit knowledge about underlying rules of the sequence learning task was assessed with three open questions at the end of Session 3 and Session 6 and a comprehensive questionnaire including additional questions at the end of the final session, Transfer 2. A questionnaire was applied in adults (see Appendix A), and a shorter version of the same questionnaire was administered for the four children, who completed the transfer session at home (see section *Study design*). In this case, children's parents were asked to pose the questions and document the answers of their children. All assessments were based on a procedure by Whitmarsh et al. (2013).

Wordings and scoring for open and specific questions from these assessments are detailed in Chapter II. From the open questions, proportions for participants reporting rule awareness vs. those who did not report rule awareness were calculated per age group and compared between groups. From the specific questions, an explicit knowledge score was calculated per age group and compared between groups (descriptive data for both open and specific questions, see Table 13).

2.3.2. Memory and language skills

To assess working memory, declarative memory and German grammar skills, we administered equivalent psychometric tests in Session 1 (Year 1) and in Session 4 (Year 2) in all age groups and normalized all test scores according to age (except for Plural German Grammar Skills in adults for which norms were not available and for which hence raw scores were analyzed).

Descriptive data for all age groups in the assessed memory skills and grammar skills, can be found in Chapter IV. This chapter contains analyses that relate AGL performance to cognitive skills in all age groups investigated in the present dissertation.

2.3.3. Additional assessments

Additional information about participants of all ages with regard to their visual and hearing development, educational background, 2nd languages and the use of (mobile) devices was collected with custom-made questionnaires at the end of Session 1. A screening tool for behavior on clinically relevant dimensions (children: CBCL, Döpfner et al., 2014; adults: BSCL, Franke, 2017) was administered in this context as well. Adult participants filled out these questionnaires themselves, while caregivers did so for participating children.

2.4. Data analysis

We characterized learning trajectories and averaged performance scores as proportion of correct test trials: Across-session learning was assessed as the mean performance of 50 test trials of each session. Within session learning trajectories were derived based on the means of 10 test trials per block.

Trials with reaction times shorter than 200 ms were disregarded, since we did not consider it feasible to successfully process the two sequences within less than this time. This exclusion criterion reduced trial numbers across all sessions by 0.36 % in 5-year olds and by 0.55% in 6-year-olds (a total of 82 excluded trials in 12 five-year-olds and in 17 six-year olds with a maximum of 12 trials excluded per subject).

To compare performance changes over time between the age groups, repeated-measures Analyses of Variance (ANOVAs) were conducted using the *ez* package in R (Lawrence, 2016), with *Age* (5-year-olds, 6-year-olds, Adults 2) as between-subject factor and *Session* (levels depending on analyses as described below and in the respective *Results* section) as within-subject factor. Here we describe on a general level, which comparisons we performed to address a certain question; the detailed analyses with all factors of the respective ANOVAs and included sample sizes are listed in the beginning of each result section.

1. For comparing start and end performance levels between Year 1 and Year 2, the following sessions were analyzed:
    - Session 4 vs. Session 1 (Start Level)
    - Session 4 vs. Session 3 (Retention)
    - Session 6 vs. Session 3 (End Level)
2. For comparing session differences between Year 1 and Year 2 (performance improvement over 3 sessions), an additional within-subject factor *Year* (Year 1 vs. Year 2) was added in the ANOVAs, resulting in the within-subject factors *Session* and *Year* with levels for *Session* described in the respective *Results* section.
3. For transfer performance relative to the first and last session with the first stimulus material in Year 2, the following sessions were analyzed:
    - Transfer 2 vs. Session 4 (Transfer Savings)
    - Transfer 2 vs. Session 6 (Transfer Loss)
4. To compare performance of the two child groups in Year 2 to age-matched naïve controls, separate ANOVAs for each group were conducted with the between-subject

factor *Group* including the levels of (1) 5-year-olds in Year 2 vs. 6-year-olds in Year 1 or (2) 6-year-olds in Year 2 vs. 7-year-olds in Year 1, respectively. ANOVAs further included the within-subject factor *Session*, with levels depending on analyses as described in the respective *Results* section.

ANOVAs were followed up with appropriate post-hoc tests. If scores were not normally distributed or had inhomogeneous variances, non-parametric tests instead of *t*-tests (Wilcoxon signed rank tests in case of paired and one-sample testing, Mann-Whitney-U tests in case of independent sample testing) were calculated. For correlation analyses of was not normally distributed data, Spearman correlation coefficients (r_s) instead of Pearson correlation coefficients (r) were calculated. Two tailed significant (<.05) *P*-values (if not indicated otherwise) were Greenhouse-Geisser-corrected (in case of violated sphericity) or Holm-corrected (in case of multiple comparisons). Effect sizes were calculated as generalized eta squared (η^2_g) for ANOVAs, as Cohen's *d* for *t*-tests and as matched rank biserial correlation (r) for Wilcoxon singed rank tests and Mann-Whitney-U-tests, respectively.

For session comparisons that involved Session 1 or Session 4, additional control analyses were conducted with proportion correct of test trials averaged over block 2 to 5 of each session (without the first block) to account for task novelty.

Due to the at home-setting for some sessions (Sessions 2, 3, 5 & 6 see Fig. 8) and complications due to the COVID-19 pandemic, there were some additional deviations from the task instructions in the final sample (2 participants with 1-2 additional task blocks completed in between sessions/at the end of one session (n = 1 adult, n = 1 five-year-old), 3 participants with 1-2 additional task blocks completed in one session, but 1-2 task blocks less in another session (n = 1 five-year-old, n = 2 six-year-olds), 1 six-year-old with a missing task block in one session, 1 six-year-old with the four sessions of Year 2 spread out across 2 weeks instead of approx. 1 week). These participants were included in the final analyses, since they did not show any systematic peculiarities in their response patterns and only the originally scheduled trials were included in case of additional task blocks completed (n = 5) or session performance was averaged over the available data in case of a missing task block (n = 5), respectively. Additionally, we checked for each of the reported analysis whether excluding these seven participants with slightly different task exposure would qualitatively change the pattern of results.

We performed equivalent Bayesian analyses for all inferential statistical analyses in the software JASP (Version 0.14.1; JASP Team, 2021), using default priors, and report the

Bayes Factor *(BF_{10})*. The *BF* helps evaluating whether the data at hand support the null-hypothesis (H_0) or the alternative hypothesis (H_1), and has been described as a suitable tool for interpreting null results (Dienes, 2014). For main and interaction effects in ANOVAs, we report the inclusion Bayes factor *(BF_{incl})* – which compares models that contain the effect of interest to equivalent models stripped of this effect – as implemented in JASP Version 0.14.1 and recommended by e.g. Mathôt (2017) and Quintana & Williams (2018). *BF* values between 1/3 and 1/10 indicate moderate evidence for the H_0, while a BF of lower than 1/10 is considered strong evidence for the H_0; a BF between 1 and 1/3 is defined as anecdotal evidence for the H_0 (Schönbrodt & Wagenmakers, 2018). *BF* values between 3 and 10 indicate moderate evidence for the H_1, while a *BF* from 10 onwards is considered as strong evidence for the H_1 and a *BF* between 1 and 3 is defined as anecdotal evidence for the H_1 (Schönbrodt & Wagenmakers, 2018). For post-hoc tests on scores that were not normally distributed, the *BF* was calculated for non-parametric test equivalents to the respective inferential tests and is reported using the default setting of data augmentation algorithms with 5 chains of 1000 iterations as implemented in JASP.

All data analyses apart from Bayesian analyses were performed in the software R (Version 4.1.0; R Core Team, 2021).

To additionally look into within-session learning, we made use of the trial-by-trial response data and fit the state-space random effects model by Smith et al. (2005) to binary responses (correct = 1, incorrect = 0) in all 150 test trials of the 3 sessions with stimulus set 1, separately for each age group and each year (Session 1 to 3 in Year 1 and Session 4 to 6 in Year 2 with the number of complete datasets per age group detailed in the above *Participants* section and in the summary tables of each respective results section below). This model estimated the trial at which learning had first occurred for the whole population (i.e., age group), by estimating an unobservable learning state process, defined as a random walk. For an estimation of the learning curves it used a state-space random effects model and Expectation-Maximization algorithm, characterizing the dynamics of the learning process as a function of trial number (Smith et al., 2005). The modeling script was provided in Matlab (Matlab, MathWorks 2020) from the website indicated by Smith et al. (2005; http://annecsmith.net/behaviorallearning.html).The estimated first learning trial from this population modeling was then used to compare within-session learning between Year 1 and Year 2 within age groups (see results section *Earlier trial-by-trial learning effects in Year 2 compared to Year 1*) and to compare within-session learning in Year 2 of 5-year-olds and 6-

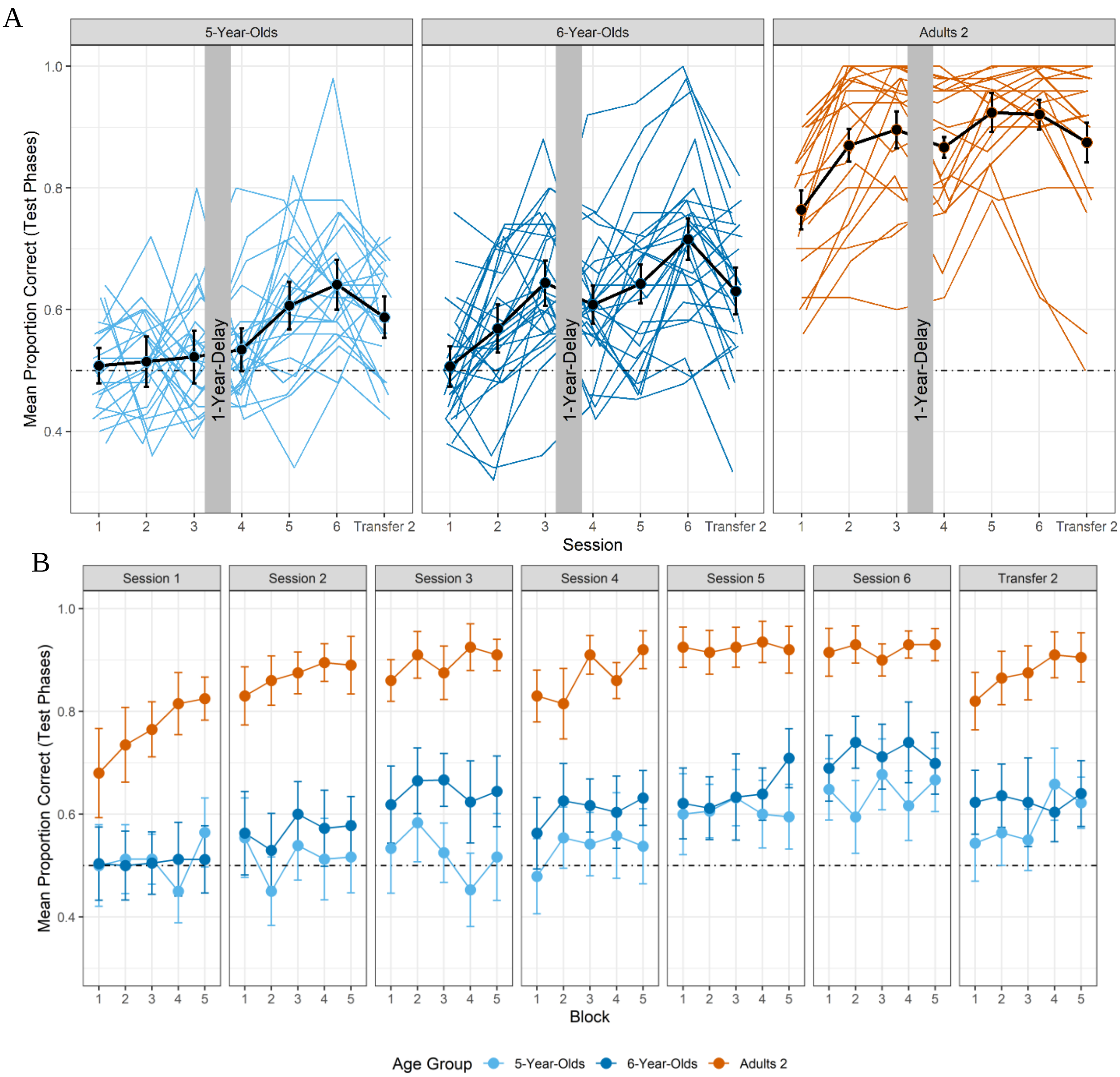
year-olds with prior learning experience to age-matched controls without prior learning experience (see results section *Earlier trial-by-trial learning effects in Year 2 due to prior in 5-year-olds and 6-year-olds compared to Controls*).

3. Results

We first tested whether all age groups performed above chance in each session, to assess whether they had learned the AG rule (Fig. 9A).

Figure 9

Performance Trajectories Across Sessions (A) and Within Sessions (B) of Year 1 & 2



Note. A: Mean proportion of correct responses in the test phases of each session for all age groups. Learning curves of single participants are depicted in color. B: Mean proportion of correct responses in the test phases of each block for all age groups. The dotted horizontal lines mark chance level performance. Error bars indicate 95% CIs corrected for within-subject comparison according to Morey (2008).

For adults, group-level performance exceeded the chance level of 0.5 (two-alternative forced-choice trials) in all sessions (all $V(20) \geq 190.00$, all $p \leq .007$, all $r \geq 0.87$, all BF_{+0} [Wilcoxon signed-rank] ≥ 6.51; one-sided).

For 6-year-olds, group-level performance exceeded the chance level after the first Session, from Session 2 in Year 1 to Transfer 2 in Year 2 (all $t(26) \geq 3.08$, all $p \leq .007$, all $r \geq 0.55$, all $BF_{+0} \geq 17.28$; one-sided). Five-year-olds did not perform above chance before Session 5 in Year 2 (Session 1 to 4: all $t(23) \leq 1.14$, all $p \geq .40$, all $r \leq 0.26$, all $BF_{+0} \leq 1.63$, one-sided; Session 5, Session 6 & Transfer 2: all $t(23) \geq 4.54$, all $p \leq .007$, all $r \geq 0.71$, all $BF_{+0} > 100$, one-sided). For all descriptive session data, see Table 6 and Appendix C (session averages without block 1). Table 6 will be referenced for all *Means(SDs)* in the following analyses to improve readability.

Table 6

A *Year 1: Proportion Correct in AGL per Session and Age Group*

	Session 1			**Session 2**			**Session 3**		
	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2
N	24	27	20	24	27	20	24	27	20
M	.51	.51	.76	.52	.57	.87	.52	.64	.90
SD	.07	.09	.11	.08	.12	.13	.10	.11	.12
Min	.38	.38	.56	.36	.32	.62	.39	.36	.60
Max	.64	.76	.92	.72	.74	1.00	.80	.88	1.00

B *Year 2: Proportion Correct in AGL per Session and Age Group*

	Session 4			**Session 5**			**Session 6**			**Transfer 2**		
	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2
N	24	27	20	24	27	20	24	27	20	24	27	20
M	.53	.61	.87	.61	.64	.92	.64	.72	.92	.59	.63	.88
SD	.09	.11	.12	.12	.13	.07	.11	.14	.11	.09	.12	.14
Min	.42	.42	.62	.34	.45	.78	.48	.48	.62	.42	.33	.50
Max	.80	.92	1.00	.82	.94	1.00	.98	1.00	1.00	.72	.82	1.00

Note. 5yo = 5-year-olds, 6yo = 6-year-olds, Ad 2 = Adults 2, Min = minimal value, Max = maximal value.

3.1. Performance in Year 2 benefits from prior learning in Year 1

3.1.1. Initial and final performance levels largely improve from Year 1 to Year 2 and are retained across a 12-month delay

We performed three ANOVAs to compare learning in Year 1 to relearning in Year 2, which are detailed in Table 7, including the addressed question, included datasets and ANOVA factors with factor levels.

Table 7

ANOVAs on Start Levels, Retention & End Levels

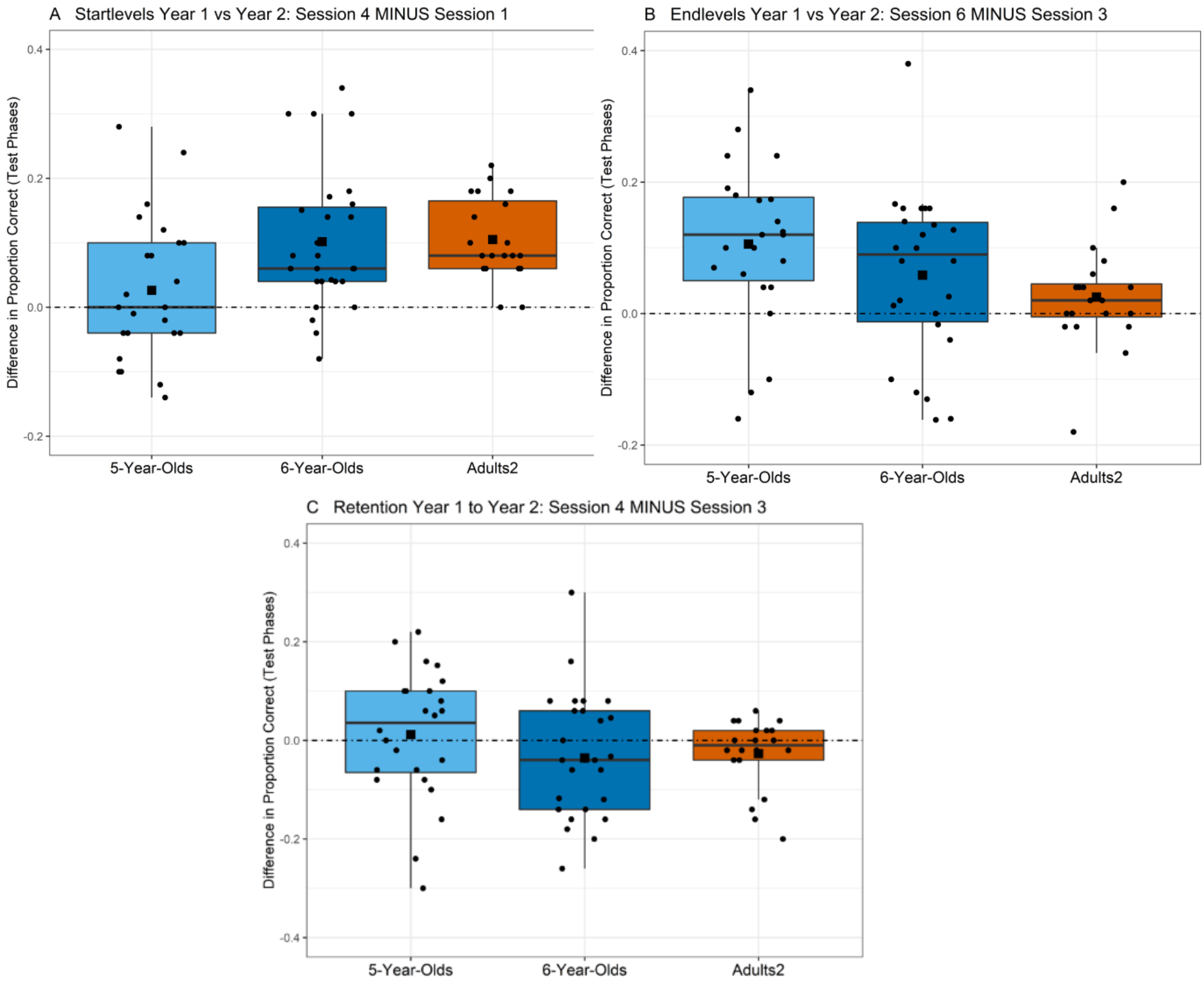
Addressed Question/ Comparison	Included datasets	Between-subject factors (levels)	Within-subject factors (levels)
Higher Start Levels in Year 2 vs. Year 1	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds,6-Year-olds, Adults 2)	Session (Session 1, Session 4)
Retention from Year 1 to Year 2	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Session (Session 3, Session 4)
Higher End Levels in Year 2 vs. Year 1	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Session (Session 3, Session 6)

First, to evaluate gains for relearning in Year 2 from prior learning experience in Year 1, we tested averaged performance in the very first session (Session 1) against averaged performance in the first session of relearning in Year 2 (Session 4) (see Fig. 10A): An ANOVA with the factors *Age* and *Session* (for factor levels, see Table 7: Higher Start Levels) revealed a significant interaction of *Age*Session* ($F(2,68) = 4.68$, $p = .011$, $\eta^2_g = .03$, $BF_{incl} = 3.67$) in addition to main effects of *Age* [3] and *Session* (both $F \geq 43.93$, both $p < .001$, both $\eta^2_g \geq .14$, both $BF_{incl} > 100$): 5-year-olds improved less than both 6-year-olds and Adults 2 from Session 1 (Year 1) to Session 4 (Year 2; group comparison Session 4 – Session 1: both $t \geq 2.45$, $p \leq .018$, $d \geq .69$, $BF_{10} \geq 3.06$; see Table 6 & Fig. 10A).

[3] Since adults continuously outperformed both child groups in all analyses of this and the next results section, differences in overall performance levels between age groups (as indicated by significant main effects of *Age*) will be pursued with post-tests only in the section that compares child groups (see *Improvement and transfer in all child groups including 7-year-olds*). I.e., 5-year-olds and 6-year-olds will be compared there. This is done to promote readability and conciseness for all results sections.

Figure 10

Session Differences of Year 1 & Year 2 for Initial & Final Performance Levels, Retention



Note. Initial performance was compared as difference in proportion correct in the test phases of Session 4 minus Session 1 (A). Final performance was compared as difference in proportion correct in the test phases of Session 6 minus Session 3 (B). Retention was compared as difference in proportion correct in the test phases of Session 4 minus Session 3 (C). Boxplots for 5-year-olds (light blue), 6-year-olds (dark blue) and Adults 2 (orange) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

Indeed, 5-year-olds did not show any significant benefit for their initial performance in Year 2 compared to Year 1 (mean difference Session 4 – Session 1 tested against 0: $t(23) = 1.16$, $p = .257$, $d = .24$, $BF_{10} = .39$). 6-year-olds and Adults 2 did not differ significantly in their performance improvement in Year 2 (group comparison Session 4 – Session 1: $U = 229.00$, $p = .38$, $r = 0.15$, $BF_{10} = .29$, see Table 6). This result pattern was confirmed in a control analysis without the first task block of each session, which yielded a significant interaction of *Age*Session* in addition to main effects of *Age* and *Session* (all $F \geq 3.43$, all $p \leq .038$, all $\eta^2_g \geq .58$, all $BF_{incl} \geq 1.44$).

Next, we tested whether all age groups had lost in performance from the last session with the first stimulus set (Session 3) in Year 1 to the first session of relearning in Year 2 (Session 4, see Fig. 10C). An ANOVA with the factors *Age* and *Session* (for factor levels, see Table 7: Retention) yielded only a significant main effect of *Age* ($F(2,68) = 88.53$, $p < .001$, $\eta^2_g = .65$, $BF_{incl} > 100$), with Adults 2 outperforming children (see Table 6, both $U < 26.00$, $p < .001$, $r = .90$, $BF_{10} > 100$) and 6-year-olds outperforming 5-year-olds (see Table 6, $t(49) = 4.33$, $p < .001$, $d = 1.22$, $BF_{10} > 100$) independent of the respective session. All age groups preserved their last performance level from Year 1 (Session 3) in the first Session in Year 2 (Session 4), with no age differences in preserved performance levels (main effect *Session* & interaction effect *Age*Session* n.s.: both $F \leq 1.48$, $p \geq .229$, $\eta^2_g = .01$, $BF_{incl} \leq .39$). The same pattern of effects emerged from a control analysis without the first task block of each session (significant main effect of *Age:* $F(2,68) = 80.91$, $p < .001$, $\eta^2_g = .63$, $BF_{incl} > 100$; all other $F \leq 1.83$, $p \geq .168$, $\eta^2_g \leq .02$, $BF_{incl} \leq .48$).

We further examined whether relearning in Year 2 resulted in a higher final performance level than in Year 1. Averaged performance in the last session with the first stimulus material was compared between Session 3 (Year 1) and Session 6 (Year 2) (difference scores for all age groups see Fig. 10B). In an ANOVA with factors *Age* and *Session* (for factor levels, see Table 7: Higher End Levels), a marginally significant interaction of *Age*Session* emerged ($F(2,68) = 3.06$, $p = .053$, $BF_{incl\,=}\,1.14$) in addition to main effects of *Age* and *Session* (both $F \geq 22.97$, $p < .001$, $\eta^2_g \geq .09$, $BF_{incl} > 100$): 5-year-olds' gain in final performance in Year 2, i.e., reaching a higher level in the last session with stimulus set 1 in Year 2 (Session 6) compared to the final session with the same stimulus set in Year 1 (Session 3), was more pronounced than the corresponding gain was in Adults 2 (see Table 6, group comparison Session 6 – Session 3: $t(42) = 2.72$, $p = .009$, $d = .82$, $BF_{10} = 5.13$). The other age groups did not differ significantly in their relearning gains for the final performance level (group comparisons Session 6 – Session 3: both $t \leq 1.24$, $p \geq .19$, $d \leq .39$, $BF_{10} \leq .60$).

An overview over the results on start levels, retention and final levels from this section is provided in Table 8. To summarize, 5-year-olds benefitted less than 6-year-olds and Adults 2 for their initial session performance in Year 2 compared to Year 1. Five-year-olds, however, showed the most pronounced gain for their final performance levels at relearning in Year 2 relative to their final performance in Year 1, which was again comparable for 6-year-olds and Adults 2. All age groups retained their final performance level from the last session in Year 1 across the one-year delay in their first session of Year 2.

Table 8

Overview of Result Patterns for Start Levels, Retention & End Levels

Addressed Question	**Session comparison**	**Main result for age difference**
Higher Start Levels (Year 2 > Year 1)	Session 4 – 1	Session 4 > Session 1 for 6-year-olds & Adults 2 → benefit in Year 2 to the same extent; Session 4 = Session 1 for 5-year-olds (no benefit)
Retention from Year 1 to Year 2	Session 4 – 3	Session 4 = Session 3 for all age groups → same degree of retention across 1 year
Higher End Levels (Year 2 > Year 1)	Session 6 – 3	Session 6 > Session 3 for all age groups, but most pronounced benefit for higher end levels in 5-year-olds; similar benefit for 6-year-olds & Adults 2

3.1.2. Performance improves across several sessions in both years and generalizes to a new stimulus set in Year 2

3.1.2.1. Improvement and transfer in 5-year-olds, 6-year-olds and Adults 2

We investigated how performance improved across three sessions with stimulus set 1 and was transferred to stimulus set 2 in three ANOVAs, which are detailed in Table 9 including the addressed question, included datasets and ANOVA factors with factor levels.

Table 9

ANOVAs on Learning Gains & Transfer Effects for Children vs. Adults 2

Addressed Question/ Comparison	**Included datasets**	**Between-subject factors (levels)**	**Within-subject factors (levels)**
Learning Gains with Stimulus Set 1 (Children vs. Adults 2)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Year (Year 1, Year 2) Session (First Session [Year 1: Session 1, Year 2: Session 4], Last Session [Year 1: Session 3, Year 2: Session 6])
Transfer Savings (Children vs. Adults 2)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Session (Transfer 2, Session 4)
Transfer Loss (Children vs. Adults 2)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Session (Transfer 2, Session 6)
Transfer compared to very first session (Children vs. Adults 2)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 20 Adults 2	Age (5-Year-olds, 6-Year-olds, Adults 2)	Session (Transfer 2, Session 1)

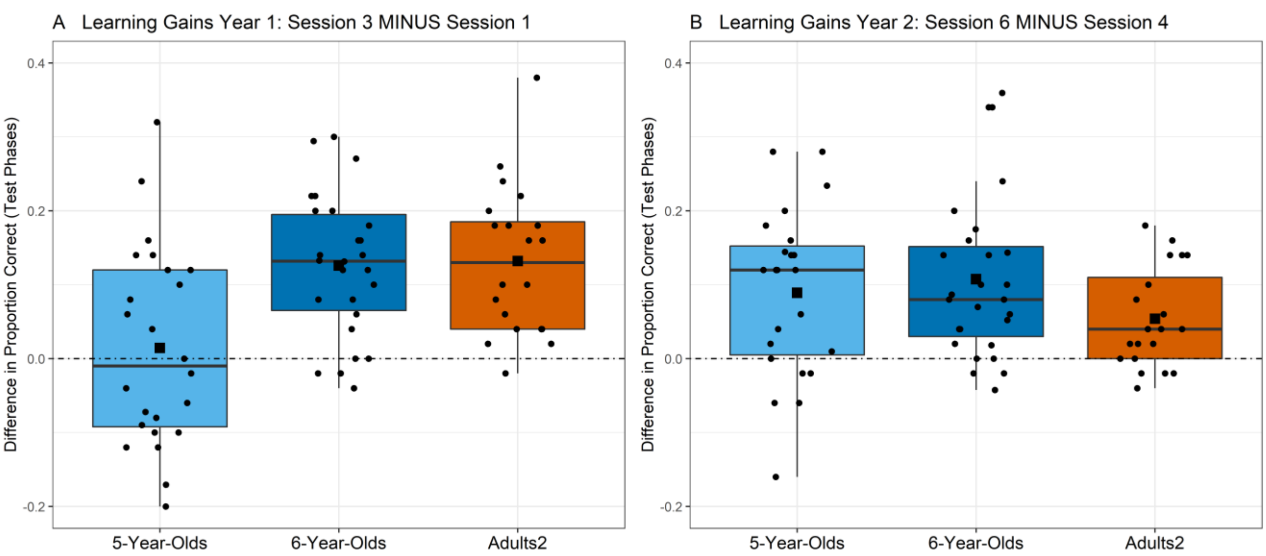
To compare performance improvements over the three sessions with the first stimulus set in both years, we tested initial learning in Year 1 (see Fig. 11A) against relearning in Year 2 (see Fig. 11B): An ANOVA with the factors *Age*, *Year* and *Session* (for factor levels, see Table 9: Learning Gains with Stimulus Set 1) yielded a significant three-way interaction, *Age*Year*Session* ($F(2,68) = 7.01$, $p = .002$, $\eta^2_g = .03$, $BF_{incl} = 21.66$), in addition to a significant two-way interaction of *Age*Session* and significant main effects of *Age*, *Year* and *Session* (all $F \geq 3.45$, $p \leq .037$, $\eta^2_g \geq .02$, $BF_{incl} > 1.39$). Post-hoc comparisons revealed that 5-year-olds improved less over three sessions in Year 1 (Session 1 to 3) than both 6-year-olds (see Table 6, group comparison Session 3 – Session 1: $t(49) = 3.56$, $p < .001$, $d = 1.0$, $BF_{10} = 36.35$) and Adults 2 (see Table 6, group comparison Session 3 – Session 1: $t(42) = 3.23$, $p = .002$, $d = .98$, $BF_{10} = 15.13$).

5-year-olds indeed did not show any performance improvement across three sessions in Year 1 (see Table 6, mean difference Session 3 – Session 1 tested against 0: $t(23) = .53$, $p = .602$, $d = .11$, $BF_{10} = .24$; no block 1: $t(23) = .35$, $p = .732$, $d = .07$, $BF_{10} = .22$). In contrast, 6-year-olds and Adults 2 did not differ in the extent to which they improved across the first three sessions of Year 1 (see Table 6, group comparison Session 3 – Session 1: $t(45) = .17$, $p = .866$, $d = .05$, $BF_{10} = .30$).

In Year 2, however, all three age groups improved to a similar degree from Session 4 to 6 (see Table 6, group comparisons Session 6 – Session 4: all $t \leq 1.53$, $p \geq .133$, $d \leq .46$, $BF_{10} \leq .76$). The above result pattern replicated in a control analysis without the first task block of each session (*Age*Year*Session*: $F(2,68) = 5.58$, $p = .006$, $\eta^2_g = .02$, $BF_{incl} = 14.43$).

Figure 11

Learning Gains in Year 1 & 2



Note. Learning Gains as difference in proportion correct responses of Session 3 and Session 1 (A: Year 1) or Session 6 and Session 4 (B: Year 2), respectively. Boxplots for 5-year-olds (light blue), 6-year-olds (dark blue) and Adults 2 (orange) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

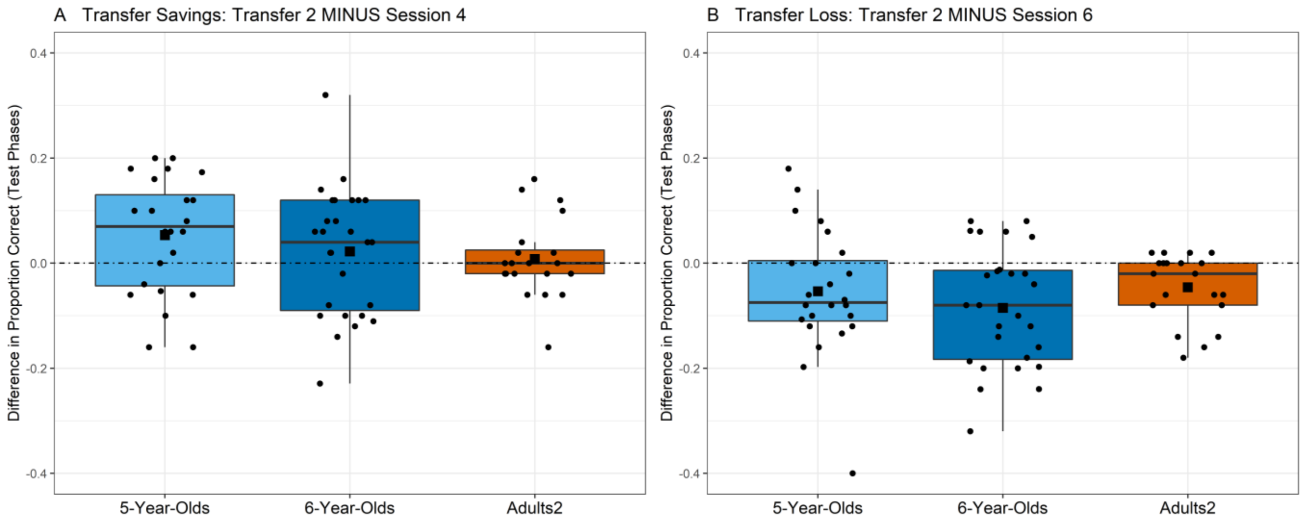
We further characterized learning across three sessions for Year 1 and Year 2 in all age groups separately by polynomial contrast analyses, which tested for linear and quadratic trends in the data (R package *emmeans*; Lenth, 2021): The increase in performance followed a linear trend in the 6-year-old and adult groups from Sessions 1 over Session 2 to Session 3 (Year 1) and from Sessions 4 over Session 5 to Session 6 (Year 2; all $p \leq .002$). Five-year-olds showed no trend for any increase across Sessions 1 to 3 (Year 1, $p = .567$), but a linear trend for the increase from Session 4 to 6 (Year 2, $p < .001$) as in 6-year-olds and Adults 2. In Adults 2, an additional quadratic trend emerged over Sessions 1 to 3 and over Sessions 4 to 6 (both $p \leq .042$).

Thus, with regard to performance increases for the first stimulus set across several sessions, 6-year-olds and Adults 2 improved to a similar degree and in a linear fashion in Year 1, while 5-year-olds showed no learning across the three sessions in Year 1. For relearning across another set of three sessions in Year 2, all age groups, including the 5-year-olds, improved to the same extent, sharing a linear performance increase. Adults' performance in both years additionally increased in an exponential fashion, reaching a very high performance level of ca. 90% accuracy from the second session of each year onwards (see Table 6).

Next, to compare how participants transferred learned regularities to a second stimulus set in Year 2 for *Transfer Savings* (see Fig. 12A), we conducted an ANOVA with the factors *Age* and *Session* (for factor levels, see Table 9: Transfer Savings).

Figure 12

Transfer Effects in Year 2



Note. Transfer Savings were calculated as difference in proportion correct in the test phases of Transfer 2 and Session 4 (A). Transfer Loss was quantified as difference in proportion correct in the test phases of Transfer 2 and Session 6 (B). Boxplots for 5-year-olds (light blue), 6-year-olds (dark blue) and Adults 2 (orange) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

In this analysis, no significant interaction of *Age**Session* (F(2,68) = 1.08, p = .347, η^2_g = .01, BF_{incl} = .27), emerged: Performance in the transfer session was better than in the first session of Year 2 (Session 4) across all age groups (significant main effects of *Age* and *Session* (both $F \geq 4.81$, $p \leq .032$, $\eta^2_g \geq .02$, $BF_{incl} \geq 1.69$), with Adults 2 outperforming both child groups (see Table 6, both $U \leq 15.00$, $p < .001$, $r \geq .84$, $BF_{10} > 100$) and 6-year-olds outperforming 5-year-olds (see Table 6, t(49) = 2.44, p = .018, d = .69, BF_{10} = 3.03) irrespective of session. This result pattern was confirmed in an analysis that excluded the first task block of both sessions. In this control analysis, the main effect of *Session*, indicating Transfer Savings in Year 2, only showed a trend towards significance (F(1,68) = 3.81, p = .055, η^2_g = 01, BF_{incl} = 1.11; *Age**Session* n.s.: F(2,68) = 0.74, p = .479, η^2_g = .01, BF_{incl} = .22). Especially Adults 2 performed on a very high level in Session 4 already (*Mean(SD)* all task blocks: .87 (.12) see Table 6, *M(SD)* without 1st task block: .88 (.11), see

Table C.2 in Appendix C), leaving little room to improve further and show any Transfer Savings (Transfer 2 – Session 4). For this reason, we checked age differences in an additional transfer comparison, detailed in the next paragraph.

Comparable savings in transfer to the new stimulus set across age replicated in an analysis that related performance in the transfer session in Year 2 (Transfer 2) to very first learning in Year 1 with the first stimulus set (Session 1). An ANOVA with the factors *Age* and *Session* (for factor levels, see Table 9: Transfer compared to very first session) revealed significant main effects of *Age* and *Session* (both $F = 62.25$, $p < .001$, $\eta^2_g \geq .21$, $BF_{incl} > 100$), but no significant interaction of *Age*Session* ($F(2,68) = 1.05$, $p = .357$, $\eta^2_g = .01$, $BF_{incl} = .26$). Overall, Adults 2 showed better performance, collapsed across Session 1 and Transfer 2, compared to both child groups (see Table 6, both $U \leq 15.00$, $p < .001$, $r \geq .90$, $BF_{10} > 100$), while 5-year-olds and 6-year-olds did not differ significantly in their overall performance average of these two sessions (see Table 6, $t(49) = 0.99$, $p = .325$, $d = .28$, $BF_{10} = .42$). This was also the case when leaving out the first task block per session, when comparing Session 1 to Transfer 2 (main effects of *Age* & *Session*: $F \geq 58.88$, $p < .001$, $\eta^2_g \geq .19$, $BF_{incl} > 100$; *Age*Session* n.s.: $F(2,68) = 0.63$, $p = .537$, $\eta^2_g < .01$, $BF_{incl} = .17$).

We further investigated preserved performance in the transfer session (Transfer 2) compared to the directly preceding session with the first stimulus set in Year 2 (i.e., Session 6) as *Transfer Loss* (see Fig. 12B). In an ANOVA with the factors *Age* and *Session* (for factor levels, see Table 9: Transfer Loss), significant main effects of *Age* and *Session* emerged (both $F(1,68) \geq 24.09$, $p < .001$, $\eta^2_g \geq .01$, $BF_{incl} > 100$; *Age*Session* n.s.: $F(2,68) = 0.96$, $p = .387$, $\eta^2_g < .01$, $BF_{incl} = .35$): All age groups performed worse in the transfer session with the new stimulus set than in the preceding relearning session with the first stimulus set. Adults 2 outperformed both child groups (see Table 6, both $U \leq 31.50$, $p < .001$, $r \geq .79$, $BF_{10} > 100$), and 6-year-olds outperformed 5-year-olds (see Table 6, $t(49) = 2.14$, $p = .037$, $d = .60$, $BF_{10} = 1.77$), irrespective of session.

Thus, 5-year-olds, 6-year-olds and Adults 2 generalized regularities from the first stimulus set to a new stimulus set to the same degree: They showed a similar performance gain in the transfer session compared to the very first learning session of Year 1 (Session 1) and compared to the first session of relearning in Year 2 (Session 4, *Transfer Savings*). All age groups lost in transfer performance relative to their final performance level reached with the first stimulus set in Year 2 (Session 6, *Transfer Loss*), and did so to the same extent independent of age. Table 11 will summarize these results for all age groups combined after the next section that additionally includes the third child group of this dissertation (7-year-olds of Project 1).

3.1.2.2. Improvement and transfer in all child groups including 7-year-olds

In addition to the three age groups considered so far (5-year-olds, 6-year-olds, Adults 2), we performed analyses on performance improvements (repeated learning with stimulus set 1) and transfer effects (generalizing the AG to stimulus set 2) on all three children groups from this dissertation, including 7-year-old children. Seven-year-olds completed an additional transfer session in Year 1 after Session 3 (see Table 1), but based on comparisons between the two adult groups included in this project (Adults 1: additional transfer in Year 1, same as 7-year-olds vs. Adults 2: no transfer in Year 1, same as 5-year-olds & 6-year-olds, see Table 1), having an additional transfer session in Year 1 seemed to have a negligible impact on relearning in Year 2 (no group differences for absolute performance levels in transfer sessions, no group differences for transfer effects relative to Session 1, 4 and 6; see Appendix C). Thus, 7-year-olds were included to provide a more comprehensive picture of developmental effects on multi-session learning and generalization in the present sequence learning paradigm. Analogous analyses to the previous sections were conducted, but this time for the three child groups (see Table 10 with the addressed question, included datasets and analysis factors with levels). Table 4 (7-year-olds) and Table 6 (5- & 6-year-olds) will be referenced for all *Means(SDs)* in the following analyses.

Table 10

ANOVAs on Learning Gains & Transfer Effects in all 3 Child Groups including 7-Year-Olds

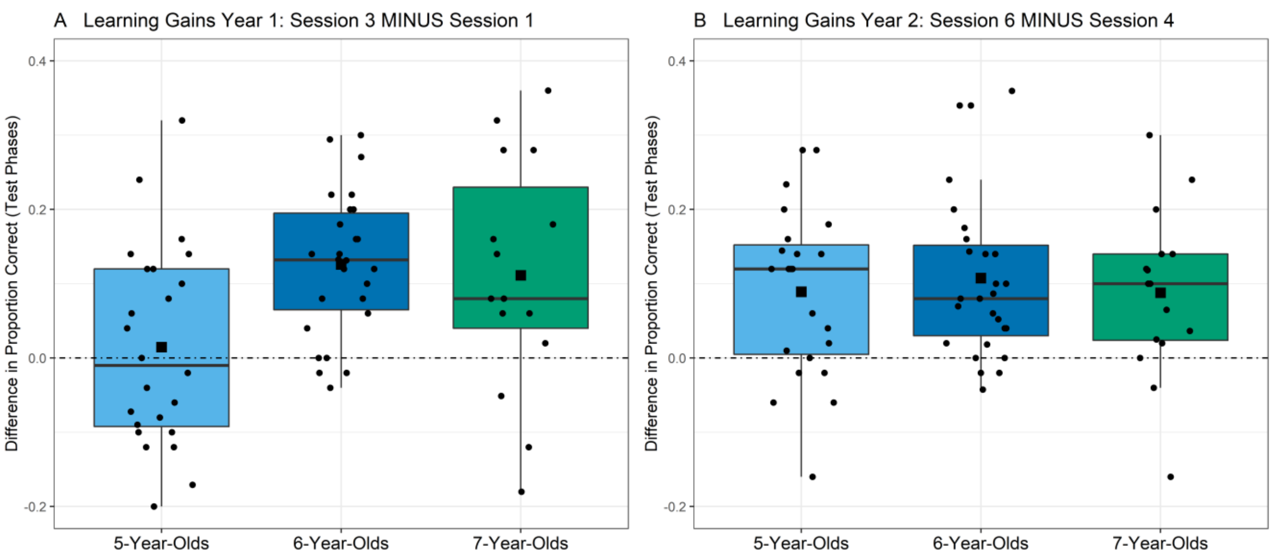
Addressed Question/ Comparison	**Included datasets**	**Between-subject factors (levels)**	**Within-subject factors (levels)**
Learning Gains with Stimulus Set 1 (3 Child Groups)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 27 7-Year-Olds	Age (5-Year-Olds, 6-Year-Olds, 7-Year-Olds)	Year (Year 1, Year 2) Session (First Session [Year 1: Session 1, Year 2: Session 4], Last Session [Year 1: Session 3, Year 2: Session 6])
Transfer Savings (3 Child Groups)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 27 7-Year-Olds	Age (5-Year-Olds, 6-Year-Olds, 7-Year-Olds)	Session (Transfer 2, Session 4)
Transfer Loss (3 Child Groups)	*N* = 24 5-Year-Olds *N* = 27 6-Year-Olds *N* = 27 7-Year-Olds	Age (5-Year-Olds, 6-Year-Olds, 7-Year-Olds)	Session (Transfer 2, Session 6)

These analyses entailed one ANOVA on performance improvements over three sessions in Year 1 vs. Year 2 (see Fig. 13) and two ANOVAs on transfer effects in Year 2 on *Transfer Savings* and *Transfer Loss* (see Fig. 14).

The ANOVA testing performance improvements over the three sessions in Year 1 vs. Year 2 included the factors *Age, Year* and *Session* (for factor levels, see Table 10: Learning Gains with Stimulus Set 1) and yielded a significant three-way interaction of *Age*Year*Session* ($F(2,64) = 4.19$, $p = .020$, $\eta^2_g = .02$, $BF_{incl} = 2.17$) in addition to significant main effects of *Age*, *Year* and *Session* (all $F \geq 65.10$, $p < .001$, $\eta^2_g \geq .16$, $BF_{incl} > 100$; same pattern of results in an analogue ANOVA without the first task block of each session with *Age*Year*Session* ($F(2,64) = 3.49$, $p = .037$, $\eta^2_g = .02$, $BF_{incl} = 1.43$)): 5-year-olds improved less than 6-year-olds and 7-year-olds in Year 1 from Session 1 to Session 3 (see Table 6, both $t \geq 2.40$, $p \leq .021$, $d \geq .78$, $BF_{10} \geq 2.80$), whereas 6-year-olds and 7-year-olds improved to a similar degree in Year 1 ($U = 221.50$, $p = .900$, $r = .03$, $BF_{10} = .31$, see Table 4 & 6 and Fig. 13A).

Figure 13

Learning Gains for all Child Groups in Year 1 & Year 2



Note. Learning Gains as difference in proportion correct responses of Session 3 and Session 1 (A: Year 1) or Session 6 and Session 4 (B: Year 2), respectively. Boxplots for 5-year-olds (light blue), 6-year-olds (dark blue) and 7-year-olds (green) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

In Year 2, all three child groups showed a comparable increase in performance from Session 4 to Session 6 (see Table 4 & 6, all $U \geq 206.50$, $p \geq .699$, $r \leq .08$, $BF_{10} \leq .35$, see Fig. 13B). Overall, older children performed on a higher level than younger children (see Table 4 & 6, pooled performance of Sessions 1, 3, 4, 6: 5-year-olds < 6-year-olds < 7-year-olds: all $t \geq 2.53$, $p \leq .015$, $d \geq .80$, $BF_{10} \geq 3.57$).

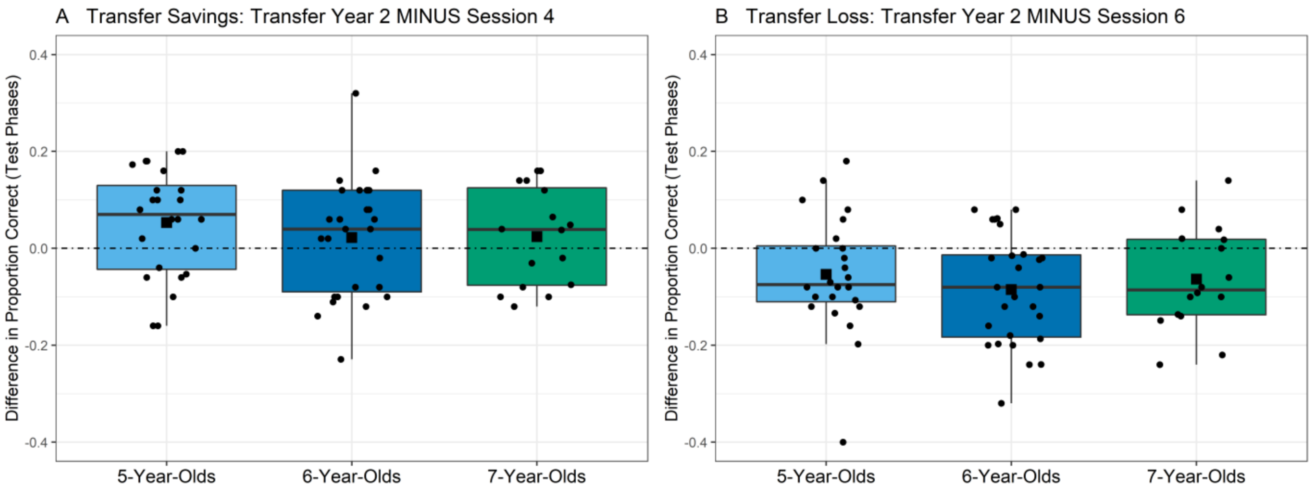
Relating this analysis to the previous section that included Adults 2, this means that from 6 years onwards, children improved to a similar degree over the three sessions with the first stimulus set in Year 1 (Session 1 to Session 3). For relearning rates with the same stimulus set across three sessions in Year 2 (Session 4 to Session 6), all age groups, from age 5 years onwards, benefitted to the same extent from their prior learning in Year 1. In their overall performance levels with the first stimulus set in these sessions, an age gradient emerged with 7-year-olds outperforming both 6-year-olds and 5-year-olds, and 6-year-olds outperforming 5-year-olds.

Investigating transfer effects in all three child groups, a first ANOVA tested how children transferred learned regularities from the first relearning session (Session 4) to a second stimulus set in Year 2 (*Transfer Savings*, for factor levels of *Age* & *Session*, see Table 10). In this analysis, no age differences in transfer emerged (see Fig. 14A) as reflected in a non-significant interaction of *Age**Session* ($F(2,64) = 0.56$, $p = .572$, $\eta^2_g < .01$, $BF_{incl} = .20$, significant main effects of *Age* & *Session*: both $F \geq 5.56$, $p \leq .021$, $\eta^2_g \geq .02$, $BF_{incl} \geq 2.60$; same pattern without first task block: *Age**Session* ($F(2,64) = 0.96$, $p = .390$, $\eta^2_g < .01$, $BF_{incl} = .24$, but no significant main effect of *Session*: $F(1,64) = 2.33$, $p = .13$, $\eta^2_g = .01$, $BF_{incl} = .68$). Overall, older children performed on a higher level than younger children (see Table 4 & 6, pooled performance of Sessions 4 & Transfer 2: 5-year-olds < 6-year-olds < 7-year-olds: all $t \geq 2.44$, $p \leq .018$, $d \geq .69$, $BF_{10} \geq 3.03$).

To test transfer as preserved performance from Session 6 to the subsequent transfer session in Year 2 (Transfer 2), we computed an ANOVA with the factors *Age* & *Session Transfer Loss*, for factor levels, see Table 10). Again, the three child groups did not differ in their amount of performance loss from stimulus set 1 to stimulus set 2 (see Fig. 14B), reflected in a non-significant interaction of *Age**Session* ($F(2,64) = 0.50$, $p = .606$, $\eta^2_g < .01$, $BF_{incl} = .18$; main effects of *Age* & *Session*: both $F \geq 10.08$, $p < .001$, $\eta^2_g \geq .07$, $BF_{incl} > 100$). Overall, older children outperformed younger children (see Table 4 & 6, pooled performance of Sessions 6 & Transfer 2: 5-year-olds < 6-year-olds < 7-year-olds: all $t \geq 2.14$, $p \leq .037$, $d \geq .60$, $BF_{10} \geq 1.77$).

Figure 14

Transfer Effects for all Child Groups in Year 2



Note. Transfer Savings were calculated as difference in proportion correct in the test phases of Transfer 2 and Session 4 (A). Transfer Loss was quantified as difference in proportion correct in the test phases of Transfer 2 and Session 6 (B). Boxplots for 5-year-olds (light blue), 6-year-olds (dark blue) and 7-year-olds (green) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions.

Thus, in sum, transfer effects showed no clear developmental trend for 5-year-olds, 6-year-olds and 7-year-olds, but older children performed on an overall higher level (7-year-olds > 6-year-olds > 5-year-olds). Age-invariant transfer effects emerged despite the fact that the oldest age group even had the advantage of an additional transfer session in Year 1 and, thus, was already familiar with the second stimulus set in the transfer session of Year 2. Nevertheless, all child groups showed a similar benefit for transferring learned regularities to a second stimulus set in Year 2.

All results for this whole section on age comparisons in performance improvements with the first stimulus set and transfer effects to a second stimulus set are summarized in Table 11 (incl. Adults 2).

Table 11

Overview of Results for Performance Improvements & Transfer Effects

Addressed Question	**Session comparison**	**Main result for age difference**
Learning Gains (Stimulus Set 1)	Session 3 – 1 (Year 1) vs. Session 6 – 4 (Year 2)	Year 1: Session 3 > Session 1 to the same degree in 6-year-olds, 7-year-olds, Adults 2; Session 3 = Session 1 in 5-year-olds Year 2: Session 6 > Session 4 to the same degree in all age groups (5-,6-,7-year-olds, Adults 2)
Transfer Savings	Transfer 2 – Session 4	Transfer > Session 4* in all age groups to the same extent; → age-independent performance savings for stimulus set 2
Transfer Loss	Transfer 2 – Session 6	Transfer < Session 6 in all age groups to the same extent; → age-independent loss in performance for stimulus set 2

Note. * n.s. when excluding first task block of both sessions, but age-independent Transfer Savings (n.s. Age*Session interaction) with a significant main effect of Session replicated for Transfer 2 (Year 2) > Session 1 (Year 1) in 5-year-olds, 6-year-olds and Adults 2.[4]

3.1.3. Earlier trial-by-trial learning effects in Year 2 compared to Year 1

For within-session learning, we fitted the state-space model of Smith et al. (2005) to compare within each age group the binary response data for the three sessions with the first stimulus set between Year 1 (Session 1 to 3) and Year 2 (Session 4 to 6). The state-space model by Smith et al. (2005) identified no trial for 5-year-olds in Year 1 (i.e., no stable above chance performance after a total of 150 test trials at the end of Session 3) and the 12th test trials (i.e., the beginning of the second task block in Session 4) at relearning in Year 2 as the timepoint at which learning first was evidenced by the model. For 6-year-olds, the 41st test trial (i.e., the end of Session 1, last task block) was identified as the first timepoint of learning in Year 1, while for their relearning in Year 2, the first timepoint of above-chance performance was identified already at the 6th test trial (i.e., in the first task block of Session 4). By contrast, Adults 2 showed evidence for learning from their very first test trial on in both years.

[4] Note that this additional transfer difference with Session 1 was not tested between 5-year-olds, 6-year-olds and 7-year-olds, since 7-year-olds had an additional transfer session in the end of Year 1. Checking for a non-significant *Age*Session* effect in the presence of a main effect of *Session* in a control analysis seemed to be of little use, as the main effect of *Session* would mean something different for 5- & 6-year-olds (1st vs. 7th session) as compared to 7-year-olds (1st vs. 8th session).

Thus, both child groups showed earlier learning effects in Year 2 compared to Year 1: This means they needed less exposure to grammatical sequences at relearning before they showed stable learning effects as identified by the state-space model. Overall, children needed more task exposure than adults before they started learning. Adults showed within-session learning effects after being exposed to a single learning phase of 18 grammatical sequences already in Year 1.

Table 12 lists all learning trials identified by the state-space model by Smith et al. (2005) for Year 1 and 2, including 7-year-olds, whose model fitting is described in detail in Chapter II. Seven-year-olds are discussed there in more detail as compared to Adults 1, who were matched in their complete study design and showed the same results as Adults 2 in their first learning trial (see Chapter II).

Table 12

Overview of First Learning Trials identified by the State-Space Model (Smith et al., 2005)

First Learning Trial (above chance performance)	**5-year-olds (*n* = 24)**	**6-year-olds (*n* = 25)**	**7-year-olds (*n* = 27/16)**	**Adults 2 (*n* = 20)**
Year 1 (Session 1 to 3)	NaN	41	30	1
Year 2 (Session 4 to 6)	12	6	1*	1

Note. NaN = no learning trial identified from model.

* in this group, only a subgroup of the original sample returned for the home follow-up in Year 2 (*n* = 16, see Chapter II).[5]

[5] Smith et al. (2005) put forward that their analysis operates by the concept of exchangeability, allowing for the estimation of a population learning curve from response data of each subject in a group that applies to every other subject in that group. Nevertheless, we checked the first learning trial in Year 1 for this returning subgroup to exclude the possibility that they were quicker learners in the first place than the children who did not take part in Year 2. The first learning trial in Year 1 identified for the returning subgroup was 80, which is even later than trial 30 that was identified for the whole sample.

3.1.4. Performance correlations with explicit sequence knowledge in Year 2

The number of children per age group, who spontaneously mentioned to have noticed some aspect related to sequence rules when asked about the AGL task and their decision strategies (open questions, see Methods, *Explicit Knowledge of Sequence Rules*), increased from Session 3 (33% of 5-year-olds, 56% of 6-year-olds), over Session 6 (75% of 5-year-olds, 70% of 6-year-olds) to the Transfer session (83% of 5-year-olds, 81% of 6-year-olds; see Table 13). At the end of the study in Year 2 (Transfer 2), the vast majority of both child groups (> 80%) thus reported to have noticed sequential regularities in the task. For Adults 2, all but one participant in Session 3 (Year 1) and every single participant in Session 6/Transfer 2 (Year 2) mentioned sequence rules in their answers to these open questions (see Table 13).

Table 13

Reported Explicit Knowledge about Sequence Rules per Age Group

Explicit knowledge		**5-Year-Olds (n = 24)**	**6-Year-Olds (n = 27)**	**Adults 2 (n = 20)**
Year 1	*N* Session 3: 0/1[a]	16/8	11/14 (2 NA)	1/19
Year 2	*N* Session 6: 0/1[a]	6/18	7/17 (2 NA)	0/20
	N Transfer 2: 0/1[a]	4/20	5/21 (1 NA)	0/20
	Score Transfer 2[b]: *M (SD)*	0.40 (0.24)*	0.40 (0.39)*	0.54 (0.22)

Note. NA = missing values for open questions, * n = 2 missing values for 5-year-olds, n = 3 missing values for 6-year-olds.
[a] 0 = sequence rules not mentioned in answers to open questions, 1 = sequence rules mentioned in answers to open questions.
[b] scores could range from -1 (no rule knowledge) to 1 (max. rule knowledge).

When statistically comparing proportions of participants with reported awareness about sequential rules in these three sessions between the age groups, 5-year-olds, 6-year-olds and adults differed significantly in Session 3 of Year 1 ($\chi^2(2) = 32.03$, $p < .001$, $BF_{10} > 100$), differed to a lower degree in Session 6 of Year 2 ($\chi^2(2) = 6.33$, $p = .042$, $BF_{10} = .20$), and did not differ significantly in Transfer 2 by the end of Year 2 ($\chi^2(2) = 2.48$, $p = .290$, $BF_{10} = .03$).

When comparing 5-year-olds, 6-year-olds and Adults 2 in their scores of reported explicit sequence knowledge in the end of Year 2 (assessed in the Transfer 2), they did not significantly differ ($F(2,63) = 1.57$, $p = .216$, $\eta^2_g = .05$, $BF_{incl} = .40$, see Table 13 for *Mean(SD)* of their scores).

Explicit knowledge scores were not significantly associated with any of the AGL performance differences from Year 2 in any age group (see Table 14: all $r_s \leq |.23|$, all $p \geq .891$). Given our small sample sizes, we additionally considered the Bayes Factor for each correlation (see Table 14), which did not support any relevant relationship between AGL performance scores in Year 2 and the level of rule knowledge reported in the end of Year 2, either (all $BF_{10} \leq .51$).

Table 14

Associations between Explicit Knowledge Levels and AGL Performance per Age Group

	Retention Year 1 to Year 2 AGL (Session 4 – Session 3)		**Learning Gains AGL** (Session 6 – Session 4)		**Transfer Savings AGL** (Transfer 2 – Session 4)		**Transfer Loss AGL** (Transfer 2 – Session 6)	
	r_s	BF_{10}	r_s	BF_{10}	r_s	BF_{10}	r_s	BF_{10}
5-Year-Olds (n = 22)	.21	(.41)	.03	(.29)	-.06	(.28)	-.23	(.51)
6-Year-Olds (n = 24)	-.05	(.41)	-.05	(.29)	-.14	(.38)	-.07	(.51)
Adults 2 (n = 20)	.08	(.30)	.04	(.29)	.22	(.47)	.21	(.40)

Note. r_s = Spearman Correlation Coefficient, BF_{10} = Bayes Factor.

3.2. Relearning advantages in Year 2 for 5-year-olds and 6-year-olds compared to naïve controls

In Year 2, task performance of the child groups (5-year-olds & 6-year-olds) could be influenced by more mature general cognitive skills like higher memory capacities. Thereby, AG acquisition could be better due to unspecific effects as compared to previous performance in Year 1, rather than due to the previous learning exposure. To account for such unspecific effects, control analyses were conducted in which each child group was compared to a group of naïve children of the same age (see Table 15).

Table 15

ANOVAs on Start Levels, Learning Gains & Transfer Effects for Child Groups and Controls

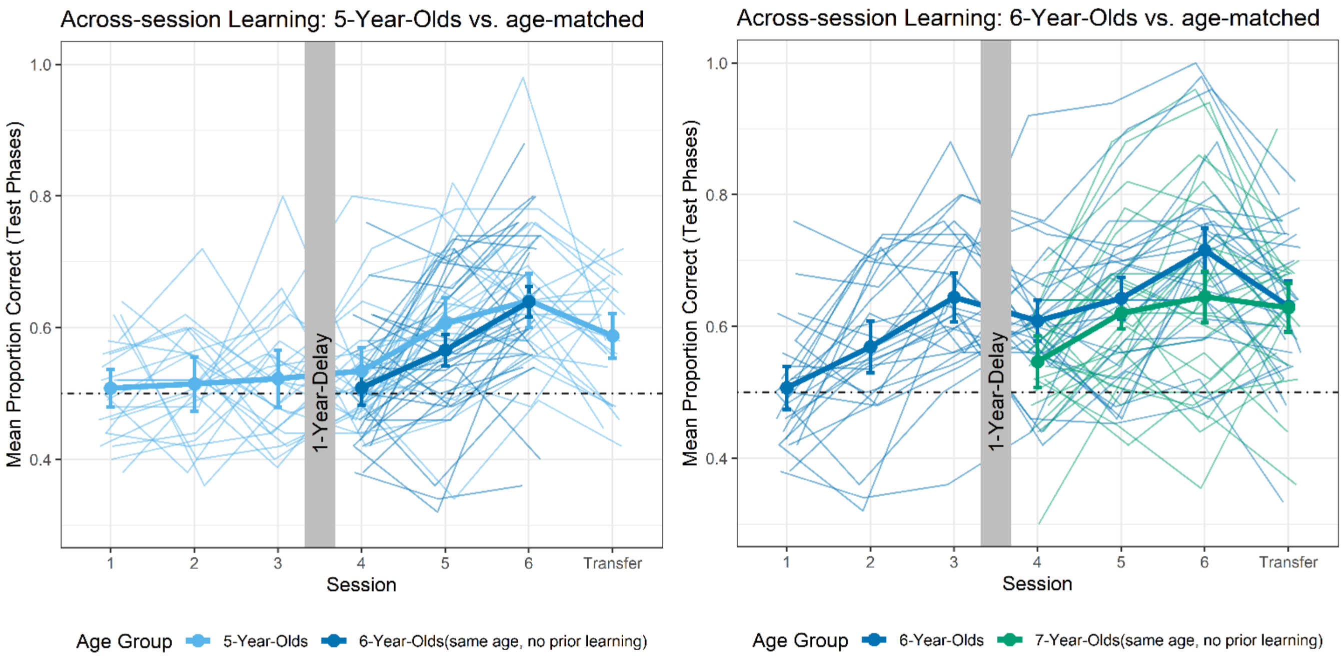
Addressed Question/ Comparison	**Included datasets**	**Between-subject factors (levels)**	**Within-subject factors (levels)**
Higher Start Levels in Year 2 (5-Year-Olds vs. Controls)	*N* = 24 5-Year-Olds Year 2 (6 years old at testing) *N* = 31 6-Year-Olds Year 1 (naïve Controls age 6)	Group for *t*-Test (5-Year-Olds Year 2, 6-Year-Olds Year 1)	Session for *t*-Test (5-Year-Olds: Session 4, 6-Year-Olds: Session 1)
Higher Start Levels in Year 2 (6-Year-Olds vs. Controls)	*N* = 27 6-Year-Olds Year 2 (7 years old at testing) *N* = 27 7-Year-Olds Year 1 (naïve Controls age 7)	Group for *t*-Test (6-Year-Olds Year 2, 7-Year-Olds Year 1)	Session for *t*-Test (6-Year-Olds: Session 4, 7-Year-Olds: Session 1)
Learning Gain with Stimulus Set 1 (5-Year-Olds vs. Controls) [Analysis 1]	*N* = 24 5-Year-Olds Year 2 (6 years old at testing) *N* = 31 6-Year-Olds Year 1 (naïve Controls age 6)	Group (5-Year-Olds Year 2, 6-Year-Olds Year 1)	Session (5-Year-Olds: Session 4 & 6, 6-Year-Olds: Session 1 & 3)
Learning Gain with Stimulus Set 1 (6-Year-Olds vs. Controls) [Analysis 2]	*N* = 27 6-Year-Olds Year 2 (7 years old at testing) *N* = 27 7-Year-Olds Year 1 (naïve Controls age 7)	Group (6-Year-Olds Year 2, 7-Year-Olds Year 1)	Session (6-Year-Olds: Session 4 & 6, 7-Year-Olds: Session 1 & 3)
Transfer Savings (6-Year-Olds vs. Controls)	*N* = 27 6-Year-Olds Year 2 (7 years old at testing) *N* = 27 7-Year-Olds Year 1 (naïve Controls age 7)	Group (6-Year-Olds Year 2, 7-Year-Olds Year 1)	Session (6-Year-Olds: Session 4 & Transfer 2, 7-Year-Olds: Session 1 & Transfer 1)
Transfer Loss (6-Year-Olds vs. Controls)	*N* = 27 6-Year-Olds Year 2 (7 years old at testing) *N* = 27 7-Year-Olds Year 1 (naïve Controls age 7)	Group (6-Year-Olds Year 2, 7-Year-Olds Year 1)	Session (6-Year-Olds: Session 6 & Transfer 2, 7-Year-Olds: Session 3 & Transfer 1)

3.2.1. Initial performance levels compared to controls

To control for maturational effects, we first compared how the group of 5-year-olds, now at age 6, performed in the first session of Year 2 (i.e., Session 4) to a group of naïve children of the same age in their very first session (6-year-olds in Year 1, see Fig. 15 left): Initial performance levels as average performance did not differ significantly between children with (5-year-olds Session 4 in Year 2) vs. without (6-year-olds Session 1 in Year 1) prior learning experience (U = 433.50, p = .30, r = .17, BF_{10} = .42; without first task block: U = 454.50, p = .162, r = .22, BF_{10} = .57). So, prior learning experience in 5-year-olds did not significantly benefit performance in the first session of Year 2, as compared to an age-matched naïve control group.

Figure 15

Performance Trajectories Across Sessions in 5-Year-Olds & 6-Year-Olds vs. age-matched Controls



Note. Mean proportion of correct responses in the test phases of each session for 5-year-olds (left) and 6-year-olds (right) with a naïve age-matched group displayed for sessions of Year 2. Learning curves of single participants are depicted in light blue (5-year-olds), dark blue (6-year-olds) and green (7-year-olds). The dotted horizontal lines mark chance level performance. Error bars indicate 95% CIs corrected for within-subject comparison according to Morey (2008).

For 6-year-olds, we compared how performance in the first session of Year 2 (Session 4), now at age 7, differed from a group of naïve children of the same age in their very first session (7-year-olds in Year 1 Session 1, see Fig. 15 right): The group of 6-year-olds outperformed the naïve control group of age-matched children in their initial performance levels, when averaged across all task blocks (U = 494.50, p = .025, r = .36, BF_{10} = 1.96)[6]. Thus, prior learning experience in 6-year-olds benefitted their performance in the initial session of Year 2 as compared to an age-matched naïve control group. When the first task block was left out, however, this performance benefit did not reach statistical significance anymore (group difference for averaged Session 4: U = 440.50, p = .112, r = .26, BF_{10} = .80).

3.2.3. Performance improvement and transfer effects compared to controls

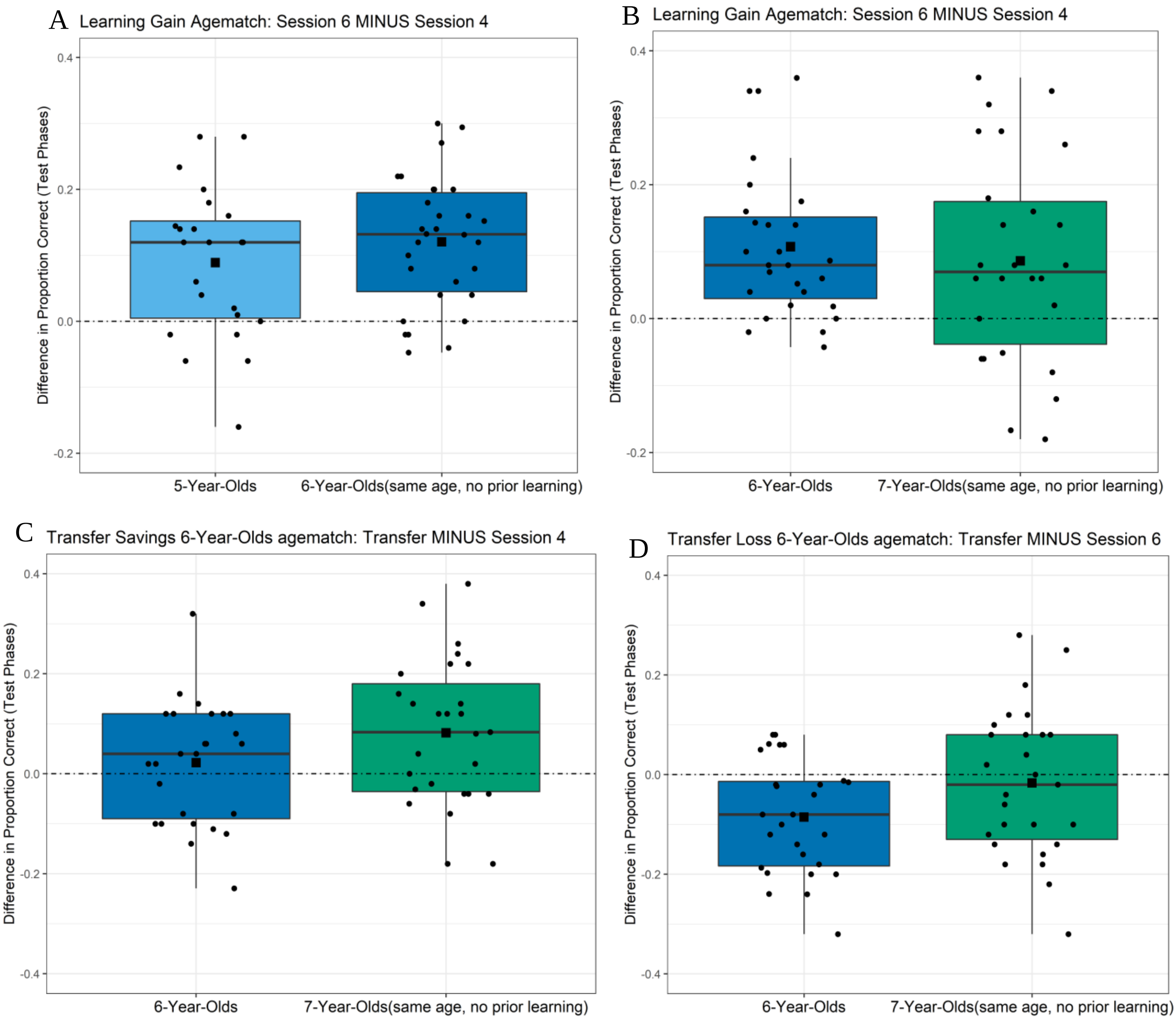
To further look into how prior learning influences performance improvement across three sessions with the first stimulus set, we also compared *Learning Gains* in Year 2 (see Fig. 16A & 16B) for 5-year-olds and 6-year-olds to an age-matched naïve control group (6-year-olds in Year 1, 7-year-olds in Year 1, respectively): Separate ANOVAs for each age group (for factor levels of *Session & Group*, see Table 15: Learning Gains Analysis 1&2) revealed no significant *Group*Session* interaction (all task blocks: both $F \leq 0.48$, $p \geq .49$, $\eta^2_g < .01$, $BF_{incl} \leq .33$,; *Group*Session* interaction without first task block: both $F \leq 1.45$, $p \geq .245$, $\eta^2_g \leq .01$, $BF_{incl} \leq .91$). In 6-year-olds, this analysis (Analysis 1) yielded a significant main effect of *Group* (all task blocks: $F(1,52) = 5.58$, $p = .022$, $\eta^2_g = .07$, $BF_{incl} = 2.67$, without first task block: $F(1,52) = 4.47$, $p = .039$, $\eta^2_g = .05$, $BF_{incl} = 1.54$), indicating significantly higher performance levels in this group for pooled performance in Sessions 4 and 6 (Year 2) than in children of the same age without prior learning experience (further statistics for *Learning Gains* ANOVAs on all task blocks: main effects of *Session* [Analysis 1&2]: both $F \geq 29.01$, $p < .001$, $\eta^2_g \geq .15$, $BF_{incl} > 100$; n.s. main effect of in 5-year-olds vs. controls [Analysis 1]:: $F(1,53) = 0.39$, $p = .533$, $\eta^2_g < .01$, $BF_{incl} = .30$).

For transfer effects, maturational influences could only be considered for the group of 6-year-olds, since for 5-year-olds there was no transfer session of an age-matched group that could be used as a reference (6-year-olds in Year 1 completed no transfer session see groups in Table 1).

[6] This performance difference was mainly driven by a performance advantage of the 6-year-olds in the "easy" test trials of Session 4 (short sequences with low ACS, see Methods section *Construction of Grammatical and Ungrammatical Sequences*), compared to naïve controls in Session 1 (group comparison for easy trials: $t(50.74) = 2.56$, $p = 0.013$, $d = .70$, $BF_{10} = 3.96$ vs. group comparison for difficult trials: $t(51.78) = 1.00$, $p = 0.324$, $d = .27$, $BF_{10} = .42$).

Figure 16

Learning Gains and Transfer Effects in 5-Year-Olds & 6-Year-Olds vs. age-matched Controls



Note. Learning Gains as difference in proportion correct responses of Session 6 and Session 4 for 5-year-olds (Year 2) or Session 3 and 1 for their control group (6-year-olds in Year 1), respectively (A). Learning Gains as difference in proportion correct responses of Session 6 and Session 4 for 6-year-olds (Year 2) or Session 3 and 1 for their control group (7-year-olds in Year 1), respectively (B). Transfer Savings as difference in proportion correct responses of Transfer 2 and Session 4 for 6-year-olds (Year 2) or Transfer 1 and Session 1 for their control group (7-year-olds in Year 1), respectively (C). Transfer Loss as difference in proportion correct responses of Transfer 2 and Session 6 for 6-year-olds (Year 2) or Transfer 1 and Session 3 for their control group (7-year-olds in Year 1), respectively (D).

We looked into transfer effects by comparing between 6-year-olds and age-matched children without prior learning experience (7-year-olds in Year 1) (1) absolute performance levels averaged in the transfer session, (2) *Transfer Savings* as transfer performance relative to performance in the first session of the same year with stimulus set 1, see Fig. 16C, and (3) *Transfer Loss* as transfer performance relative to performance in the last session of the same year with stimulus set 1, see Table 15 and Fig. 16D.

Performance in the transfer session was indistinguishable between 6-year-olds and age-matched controls without prior learning ($t(51) = 0.34$, $p = .73$, $d = .09$, $BF_{10} = .29$). *Transfer Savings* were compared in an ANOVA with the factors *Group* (between-subject) and *Session* (within-subject; for factor levels see Table 15), which revealed no significant interaction of *Group*Session* ($F(1,52) = 2.72$, $p = .105$, $\eta^2_g = .02$, $BF_{incl} = .97$; n.s. main effect of *Group*: $F(1,52) = 2.09$, $p = .155$, $\eta^2_g = .02$, $BF_{incl} = .57$, main effect of *Session* $F(1,52) = 8.27$, $p = .006$, $\eta^2_g = .06$, $BF_{incl} = 8.71$; without first task block: *Group*Session:* $F(1,52) = 2.44$, $p = .124$, $\eta^2_g = .02$, $BF_{incl} = .32$, n.s. main effect of *Group*: $F(1,52) = 0.61$, $p = .438$, $\eta^2_g = .01$, $BF_{incl} = 1.60$, main effect of *Session:* $F(1,52) = 6.25$, $p = .016$, $\eta^2_g = .04$, $BF_{incl} > 100$). *Transfer Loss* as tested in an ANOVA with factors *Group* (between-subject) and *Session* (within-subject; for factor levels see Table 15), showed a trend for an interaction of *Group*Session* ($F(1,52) = 3.69$, $p = .060$, $\eta^2_g = .02$, $BF_{incl} = 1.26$, n.s. main effect of *Group*: $F(1,52) = 1.43$, $p = .238$, $\eta^2_g = .02$, $BF_{incl} = .54$, main effect of *Session*: both $F(1,52) = 8.14$, $p = .006$, $\eta^2_g = .04$, $BF_{incl} = 5.34$). This trend was driven by 6-year-olds reaching marginally higher final levels with the first stimulus set in Session 6 than age-matched controls ($t(52) = 1.77$, $p = .082$, $d = .48$, $BF_{10} = 1.00$), while transfer in absolute performance levels was comparable in both groups (see above).

So, both child groups at age 6 and 7 years with (Year 2 for groups of 5-year-olds and 6-year-olds) and without (age-matched controls) learning experience 12 months prior to testing improved to a similar degree across three sessions with the first stimulus set. At age 7, children with prior learning experience (6-year-olds in Year 2) subsequently transferred regularities to the same extent to a new stimulus set as naïve children of the same age.

Summarizing all analyses from this section, which controlled for maturational effects on relearning in Year 2 on a session level, maturation could account for the across-session improvement in the groups of 5-year-olds and 6-year-olds, and for transfer effects in 6-year-olds. Nevertheless, these analyses provide first evidence that 6-year-olds in Year 2, now at age 7, benefitted from prior learning compared to a naïve group of the same age: 6-year-olds in Year 2 showed better performance in their initial session than naïve controls when considering all task blocks (Session 4) and performed on an overall higher level, when looking at the pooled performance for their initial and their final session with the first stimulus set (Session 4 & 6).

3.2.4. Earlier trial-by-trial learning effects in Year 2 due to prior learning in 5-year-olds and 6-year-olds compared to controls

Next, we aimed to zoom in on within-session learning in Year 2 and tested if prior learning enabled earlier learning on a trial-by-trial basis than when being first exposed to the AG. We again compared child groups of the same age who only differed in their prior learning experience in Year 1 (naïve vs. three sessions of learning with stimulus set 1). The state-space model by Smith et al. (2005) identified earlier relearning trials in the first session of Year 2 for both 5-year-olds (learning trial: 12th test trial vs. 41st test trials in age-matched control group of 6-year-olds in Year 1) and 6-year-olds (learning trial: 6th test trial vs. 30th test trial in age-matched control group of 7-year-olds in Year 1), who had completed three sessions of sequence learning one year before. Hence, on average, both child groups showed learning effects at least two task blocks earlier than naïve children of the same age. This means that they "saved" exposure to at least 36 grammatical sequences that they needed less to exhibit sequence (re)learning effects, due to their prior learning experience.

Relearning results from this whole section are summarized in Table 16 for the groups of 5-year-olds and 6-year-olds, testing their performance against an age-matched age group without prior learning experience 12 months before. Results are displayed for analyses both on a session level and on trial-by-trial performance.

Table 16

Overview of Relearning Results in Year 2 for 5-Year-Olds & 6-Year-Olds vs. naïve Controls

(Re)Learning Measure	5-Year-Olds vs. naïve Controls	6-Year-Olds vs. naïve Controls
1st Session	5yo = naïve	6yo > naïve
Learning Gains (Stimulus Set 1)	5yo = naïve	6yo = naïve (but pooled performance Session 4+6: 6yo > naïve)
Transfer Session	N.A.	6yo = naïve
Transfer – 1st Session (Transfer Savings)	N.A.	6yo = naïve
Transfer – 3rd Session (Transfer Loss)	N.A.	Trend for Loss 6yo > naïve
First Learning Trial (State-Space Model)	Earlier than naïve (trial 12 vs. trial 41)	Earlier than naïve (trial 6 vs. trial 30)

Note. 5yo = 5-year-olds, 6yo = 6-year-olds, N.A. = not applicable due to study design (no session reference with naïve age-matched group available).

4. Discussion

In Project 2, it was tested whether an earlier developmental timing of several instances of sequence learning facilitates learning rates, retention, and transfer of acquired regularities as defined by an AG in 5-year-old children, 6-year-old children and adults (Adults 2, see Table 1). All three age groups completed a visual AGL task with two sets of 7 sessions in total. These two sets of sessions were separated by a 12-month-gap after the third session, and included a transfer session in the very end that used new visual stimuli governed by the same AG (see Figure 1). For relearning after the delay, we controlled for maturational effects in the child groups by comparing them to naïve controls of the same age.

Table 17 summarizes the main findings, relating them to our hypotheses and the investigated age groups. We observed successful learning of the AG in children from age 6 years onwards, i.e., for 6-year-olds in both years and for 5-year-olds in Year 2, and in adults. Adults overall outperformed child groups and displayed a steeper learning curve, thus, showing learning at an earlier timepoint with less task exposure needed. Despite failing to perform above-chance in all three sessions of Year 1, the group of 5-year-olds improved across the three sessions after the 12-month delay (Year 2) to the same degree as the groups of 6-year-olds and adults in their relearning of Year 2. In these two older age groups, retained rule knowledge over a period of one year was evidenced by starting out at the final level reached one year earlier. Control analyses in Year 2 for the groups of 5-year-olds and 6-year-olds vs. their respective naïve age-matched controls revealed evidence for a genuine effect of prior learning on their performance after a one-year delay, rather than unspecific (maturational) effects. This was mainly evident in earlier learning effects in the first session of Year 2, modeled on a trial basis, for the groups of 5-year-olds and 6-year-olds as compared to children of the same age but without prior task exposure one year before. Notably, in this trial-based control measure of relearning in Year 2, prior task exposure also caused earlier learning effects in the group of 5-year-olds, who had failed to show learning in any behavioral markers of Year 1. An earlier developmental timing was not found to be of advantage for any of these relearning effects, however. Besides, we were not able to confirm what we had predicted from stronger (over)generalization early in development: No age differences between all available age groups emerged for transferring the learned rule knowledge to a new stimulus set at the end of the second set of sessions in Year 2, including 7-year-olds from Chapter II.

Table 17

Main results and conclusions with respective hypotheses and investigated age groups

Hypothesis	Age Groups	Main Result / Conclusion
Similar learning rate (Year 1) across several sessions independent of age	5-year-olds, 6-year-olds, Adults 2 (7-year-olds)	Supported, but no learning in 5-year-olds; Adult-like learning efficiency from 6 years onwards for using acquired rule knowledge for subsequent learning instances over 1 week
Better long-term consolidation early in development (from Year 1 to Year 2)	5-year-olds, 6-year-olds, Adults 2	No; Retained long-term rule knowledge in all age groups (not meaningful for 5-year-olds)
Quicker relearning rate (Year 2) across several sessions early in development	5-year-olds, 6-year-olds, Adults 2 (7-year-olds)	No; Similar relearning gains in all age-groups for using acquired rule knowledge after a long-term delay for subsequent learning instances
Effects of prior learning > maturational effects in children (Year 2)	5-year-olds, 6-year-olds, naïve age-matched controls	Partly supported; Preliminary support for genuine effects of prior learning, mainly as earlier within-session effects; No evidence of an earlier developmental timing being of advantage for these effects
Larger transfer to new stimuli early in development (Year 2)	5-year-olds, 6-year-olds, Adults 2 (7-year-olds)	No; Similar transfer of acquired rule knowledge to new set of visual stimuli in all age-groups

Emerging awareness about sequence rules was assessed in open questions about the task at three points of time throughout the study (at the end of Session 3, Session 6 & Transfer 2). In their responses, a higher proportion of 6-year-olds and adults compared to 5-year-olds reported to have noticed sequence rules in the AGL task already in Year 1 before the delay – an age pattern that ameliorated in Year 2 after the delay. Levels of quantified rule knowledge in the end of Year 2, about legal item positions and item-item transitions, did not differ between 5-year-olds, 6-year-olds and adults. These knowledge levels were not significantly associated with Year 2 AGL performance parameters in any age group.

I will go through these findings one by one and relate them to the existing literature in the following paragraphs. This discussion will mainly focus on developmental changes in learning mechanisms, the concept of “savings” from prior learning, and the question which factors might influence the degree to which young children are able to generalize their rule knowledge. A broader discussion on how findings of this chapter speak to more general concepts like plasticity processes and, relatedly, sensitive periods in development, follows in Chapter V. This combined discussion in the end happens in an effort to combine results from all projects of the present dissertation to derive conceptual implications.

4.1. Developmental changes in learning mechanisms show on different timescales

In Year 1, children age 6 years and adults successfully learned the visually presented AG, with adults outperforming children but similar learning gains across three sessions in both age groups. This finding extends previous multi-session studies on children age 8-12 years (Ferman & Karni, 2010; Smalle, Page, et al., 2017), which showed that auditory sequence rules are acquired with similar learning rates, but are applied with an overall higher accuracy in adults than in children. Based on our data, these age trends can be extended to the learning of younger children (age 6 years) in the visual domain with complex sequential regularities as defined by an AG. Our results on overall performance levels of 5-, 6-, 7-year-olds and adults are furthermore in line with previous studies that indicated better performance in visual sequence learning with age, across the range from 5 to 12 years (Arciuli & Simpson, 2011; Raviv & Arnon, 2017; Shufaniya & Arnon, 2018) in childhood. This improved performance across development was extended into adulthood in a sample of age 6 to 30 years by Schlichting et al. (2017). At the same time, previous studies found that despite age differences in overall performance levels, the relative amount of short-term retention (across 24 hours) was indistinguishable between age groups (Juhász & Németh, 2018; Tóth-Fáber et al., 2023). This aligns well with the indistinguishable across-session trajectories we observed for 6-year-olds and adults, substantiating a similar benefit in children and adults from previous learning across short-term delays of one to several nights.

Five-year-old children failed to reach above-chance performance in all three sessions of Year 1 (Session 1 to 3), which corroborates reports of chance performance in young age groups from previous studies on single-session sequence learning: After exposure to visual and auditory triplets, 5- to 6-year-olds’ performance did not exceed chance in the study of Raviv and Arnon (2017). Even older children, 6 to 9 years old (mean age of 8.47 years), did not perform better than chance in an auditory AGL task (Pavlidou & Bogaerts 2019). Thus, the question arises of what determines whether young children display behavioral effects in sequence learning? For linguistic stimuli (preferentially syllables in the auditory domain; Nowak & Baggio, 2017) and less complex regularities (i.e., single transitions between 2 adjacent items or easy to discriminate foils at test; Forest et al., 2021; Witt et al., 2013; Witt & Vinter, 2017), also children of age 5 years have been observed to show learning effects in behavior (for a discussion on the special case of learning linguistic material see Forest et al., 2023). A very high amount of exposure to sequential regularities, however, does not seem to guarantee successful learning reflected in behavior (see also chance performance persisting for new items in 8-year-olds in the auditory domain after 5 additional sessions in Ferman & Karni, 2010): In our study, we exposed all participants to abundant rule-abiding input (exposure to 270 grammatical sequences in total) and provided performance feedback (for 150 test trials in total) across three sessions of visual AGL in Year 1. Nevertheless, we did not observe stable above-chance performance in 5-year-olds.

A genuine age-dependent improvement in learning capabilities might thus account for differing AGL performance in our groups of 5-year-olds and 6-year-olds in Year 1. So, what changes in middle childhood to enable more stable behavioral effects in these direct markers of sequence learning (Forest et al., 2023) at the age of approximately 6 years? Middle childhood has been described as a transition time that constitutes a developmental switch point (Del Giudice, 2014), entailing a shift in cognitive and corresponding neural mechanisms (Bunge & Zelazo, 2016; M. H. Johnson & Munakata, 2005; Ramscar & Gitcho, 2007). This more mature cognitive skill set, likely amplified by external events like entering (pre)school (Brod et al., 2017), could cause children from 6 years on to be better equipped on multiple levels for succeeding in our AGL task: These levels probably include attending to relevant task features like legal sequence transitions, suppressing prepotent responses when asked to choose the legal sequence out of two presented alternatives, and adapting behavior to the performance feedback. Related to such aspects of stronger behavioral control with age, underlying neurocognitive mechanisms might already start to shift towards more explicit processing and better accessible, more comprehensive representations of environmental regularities around age 6 (shift from childhood to adulthood reviewed recently in Conway, 2020). In support of a respective age trend, Witt and Vinter (2012) characterized single-session AGL in the same age range as our study (four age groups: 5-, 6-, 7- and 8-year-olds), by measuring which type of sequences children built in a self-generation task after exposure to visual sequences that adhered to an AG. The authors compared produced sequence patterns of all child groups to theoretical profiles predicted from stimulus-specific, mild rule-based and strong rule-based accounts of AGL. They report that produced sequence patterns reflecting the acquisition of underlying rules (as opposed to such reflecting solely the reproduction of surface features) first emerge at age 6 years. Chapter V will point out that this shift in learning around age 6 years has been observed in neighboring domains as well, like perceptual learning.

A change in the predominantly recruited learning mechanisms from earlier childhood to adulthood could be reflected in our measures of within-session learning as well: A descriptive age-gradient emerged for earlier within-session learning in older children and adults (first learning trial from state-space model by Smith et al., 2005 in Year 1: trial 1 in Adults 2 < trial 30 in 7-year-olds < trial 41 in 6-year-olds). Adults displayed a steeper learning curve than children with evidence for above-chance discrimination performance between legal and illegal sequences after as little input as 18 legal sequences for exposure (i.e., in their very first test phase after a single learning phase). Janacsek et al. (2012) (further elaborated by Daltrozzo & Conway, 2014; Nemeth et al., 2013) suggested that across development, an explicit learning or “model-based” system takes over, which relies more on attentional resources, behavioral control and prior knowledge than the “model-free” system prevailing earlier in childhood. These more supervised learning mechanisms governing sequence learning in adults could be reflected in their quick acquisition rate followed by a long asymptote – opposed to children, whose performance increased in a rather linear fashion. However, we cannot exclude the possibility that learning in adults would look different with more challenging stimulus material that prevents ceiling effects, e.g., a more complex AG or longer sequences (> 7 items). Since our main aim was to compare meaningful (above chance) trajectories between child groups, we chose AG complexity (according to Schiff & Katan, 2014) and the maximum sequence length (similar to Nowak & Baggio, 2017; Witt & Vinter, 2012) in a way that promised to be suitable for children of age 5 to 7 years[7].

4.2. Rule knowledge facilitates relearning performance after a one-year delay

After the one-year delay, all groups started out at their final level reached in Year 1, indicating retained rule knowledge across this extended time period in the groups of 6-year-olds and adults. This is consistent with previous findings on long-term retention of sequence knowledge across two months (Ferman & Karni, 2010) and one year (Smalle, Page, et al.,

[7] All stimulus material and procedures of the current study were piloted in a different sample of twenty-seven 5-7-year-olds (mean age 73 months, range 61-85 months, 11 female), to confirm successful learning of 3-to-7-item sequences and check that the applied stimulus sets (animal cars making up trains, adapted from Rosas et al. (2010), and color segments making up flags, adapted from Witt and Vinter (2012) are learned equally well by children of this age.

2017) for single auditory rules in older children (8-12 years old). By directly comparing children with adults and providing evidence for adult-like retention in children at 6 years of age (at first exposure) already, we additionally extend separate investigations of visuomotor retention across a one year delay on 9-15-year-old's (Tóth-Fáber et al., 2021) and adults' (Kóbor et al., 2017) learning. The current study differed from previous studies by measuring remaining rule knowledge for *quicker and additive relearning* after a delay (including re-exposure to the AG), rather than mere *retention*. The results nevertheless corroborate findings of successful retention in children and adults in this slightly different measure after a long-term delay, when being re-exposed to the familiar sequence rules.

Despite failing to perform above-chance in all three sessions of Year 1, the group of 5-year-olds improved across three more learning sessions after the 12-month delay (Session 4 to 6 in Year 2) to the same extent as the groups of 6-year-olds and adults. We will discuss possible maturational influences on these relearning effects in children later (see *Long-term benefits from prior learning persist in children after controlling for unspecific maturational effects).* First, this finding of age-invariant relearning rates adds to available literature on retention by characterizing how previous rule knowledge is used in multiple additional sessions after a long-term delay, instead of in a single follow-up (Ferman & Karni, 2010, 2014; Kóbor et al., 2017; Smalle, Page, et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021). We find that, on a session level, all age groups showed (additional) performance gains from three more learning sessions in Year 2. This relearning lead to better rule knowledge in the end of this second set of sessions as compared to each group's final performance level in Year 1. This corroborates the literature on age-independent retention up to one year in auditory (Ferman & Karni, 2010; Smalle, Page, et al., 2017) und visuo-motor (Kóbor et al., 2017; Tóth-Fáber, Janacsek, & Németh, 2021) sequence learning for children age 8-15 years and adults, as mentioned above. Adding evidence from our multi-session relearning measure to these reported retention effects, our results suggest that already children at age 7 years (6-year-olds in Year 2) are able to draw on their rule knowledge acquired several months ago and use it for further learning instances with adult-like efficiency. How this finding informs neurocognitive models of sequence learning and plasticity processes, which might underly the observed behavior, will be discussed in Chapter V. One limitation for this child-adult comparison, however, are very high performance levels in adults, who might have improved to a greater degree in a more difficult task (as discussed in detail above).

Despite adult-like relearning rates of sequential regularities in all our age groups on a session level, older age seemed to drive earlier within-session relearning. This was reflected in relearning modeled on a trial basis after the delay, showing that performance in older children and adults numerically exceeded the chance level at an earlier timepoint in the first session of Year 2 (first learning trial from state-space model by Smith et al., 2005 in Year 2: trial 1 in 7-year-olds & Adults 2 < trial 6 in 6-year-olds < trial 12 in 5-year-olds). From age 8 years (7-year-olds in Year 2) on, however, children displayed these within-session relearning effects from trial-based modeling as early as adults, i.e., in the very first test trial after a single learning phase with exposure to only 18 grammatical sequences. These findings imply that a shift towards more explicit learning mechanisms across development as discussed above (Conway, 2020; Daltrozzo & Conway, 2014; Janacsek et al., 2012; Nemeth et al., 2013) does not only affect the acquisition of sequential regularities at a single encounter. Additionally, this shift might result in better access to and retrieval of previously acquired representations of such regularities across an extended time period. Relatedly, older children and adults might consolidate their (previous) learning in more explicit memory traces (also discussed in H. Liu et al., 2023). Explicit memory traces have been found to become more abstract with time and more malleable to later learning than implicit memory traces (H. Liu et al., 2023), possibly putting older children and adults at an advantage in our multi-session setting with complex underlying sequence rules.

Interpreting age differences in within-session learning effects here warrants caution, however, given that modeling results only varied slightly in the investigated age groups (1^{st} learning trial in Year 2 identified after the first learning phase for 6-year-olds, 7-year-olds & Adults 2; for 5-year-olds after the second learning phase) and age differences therein were not statistically tested.

4.3. Long-term benefits from prior learning persist in children after controlling for unspecific maturational effects

Since children grew one year older over the course of our study, we wondered to what extent their matured cognitive skill set relative to their prior experience with the AGL task before the delay benefitted task performance in Year 2. Addressing this question, a contrast of 5-year-olds and 6-year-olds in Year 2 with naïve children of the same age provides first support in this age range for a genuine effect of prior learning on task performance after the one-year delay, after controlling for general maturation. The effect of prior learning was expressed on a session level in the group of 6-year-olds, in that they applied sequence rules better in the initial session of Year 2 (Session 4)[8] and displayed overall higher levels of rule knowledge in the first and last session of Year 2 (pooled performance of Session 4 & 6), both compared to age-matched controls. This experience-dependent benefit became even more clear from modeling performance on a trial basis (state-space model by Smith et al., 2005), which implied that the groups of 5-year-olds and 6-year-olds in Year 2 performed above chance more than two task blocks earlier compared to children of the same age, who had not been exposed to the AGL task one year before (see Table 1). Strikingly, earlier relearning in Year 2 was reflected in this within-session measure for both child groups – including the group of 5-year-olds, who had failed to show any learning effects in Year 1. Given their chance performance as a group in all three sessions of Year 1, it seems surprising that our youngest age group was able to draw on their previous exposure to visual regularities when reencountering the same task one year later.

One concept that tries to capture this phenomenon of relearning advantages are “savings” from previous experiences (less time or effort needed at relearning due to prior learning of the same material; Ebbinghaus, 1880). Hübener and Bonhoeffer (2014) argue that prior learning leads to enhanced plasticity, with an underlying infrastructure that might stay temporarily dormant or “hidden” until being re-exposed to the same learning material. Older studies on behavioral savings from other task domains reported different age trends for children in the age range of the present study vs. adults: Livosky and Sugar (1992) observed the greatest advantage in 3-year-olds (> 5-year-olds > young adults) in learning familiar relative to new picture pairs after two weeks’ time. Parkin and Streete (1988) found comparable relearning benefits for 3-year-olds, 5-year-olds, 7-year-olds and adults in a recognition task with fragmented pictures after one-hour and after a two-week delay. Our results extend these findings by showing that behavioral savings emerge across extended time periods (up to one year) in visual sequence learning of adults and children of age 5 and 6 years. Savings from prior learning in our group of 5-year-olds furthermore stress that prior learning does not need to be reflected in behavior in order to leave memory traces that can be reactivated at a later time, supporting behavioral advantages in the long run (for possible mechanisms of structural plasticity underlying this relearning behavior, see Hofer &

[8] However, this performance advantage did not remain statistically significant, when analyzed without the first task block to account for 6-year-olds greater general task familiarity as compared to controls.

Bonhoeffer, 2010; Xu et al., 2009; Yang et al., 2009, discussed in Chapter V). Our data, however, provide no evidence for more effective savings from an earlier developmental timing of learning experiences (first exposure with age 5 years vs. 6 years of age). This corroborates what Parkin and Streete (1988) observed for 3-year-olds, 5-year-olds, 7-year-olds and adults across shorter timescales in a different task.

I will discuss later, how savings in learning observed in the present study might speak to underlying neural adaptations in response to prior experiences, and to the related concept of sensitive periods in development (Chapter V).

4.4. Rule knowledge generalizes to new surface features after a one-year delay independent of age

Contrary to our expectations based on stronger (over)generalization early in development (Keresztes et al., 2018; Marcus et al., 1992; Ngo et al., 2018), all age groups of the present study (5-year-olds, 6-year-olds, Adults 2 , and, additionally included 7-year-olds from Project 1) transferred the learned rule knowledge to the same degree to a new stimulus set at the end of the second set of sessions in Year 2.

This is in line with single-session studies, which reported that adults and children aged 3-6 years (Nowak & Baggio, 2017) or 6-9 years (Jung et al., 2020) respectively, generalized learned regularities to new items in the auditory domain (Nowak & Baggio, 2017) and to new instances of underlying categories in the visual domain (Jung et al., 2020). Our findings challenge the proposition that children younger than 12 years of age can only apply a learned rule to new items in a multi-session setting, if this rule has been explicitly instructed in the beginning (Ferman & Karni, 2010, 2014). Instead, we show that two sets of three sessions each, spread out across a one-year delay, suffice to subsequently enable successful transfer of complex visual regularities. This was demonstrated in children who were as young as 5 years old at first exposure and without explicitly verbalizing the used sequence rules at any time throughout the study.

Hence, one might wonder which conditions facilitate the generalization of regularities across multiple learning instances in (young) children? Several authors propose that offline periods in general promote the extraction and representation of underlying regularities by replay-induced strengthening of memory representations (Lerner & Gluck, 2019; Liu et al., 2019; Wilhelm et al., 2012). They argue that this mechanism might underlie transfer effects in learning at the level of cortical circuits. This proposition fits well with the “spacing effect” from category and concept learning in children (Vlach, 2014), which has been recently applied to statistical learning (Forest et al., 2023). This approach suggests that forgetting which happens between learning experiences supports generalization, as shared features of input (in our case, underlying sequential regularities) are reactivated more frequently and, consequently, are less likely to be forgotten than specific information in the same input (in our case, surface features of stimulus sets; “forgetting-as-abstraction” by Vlach, 2014). Young children were shown to have high forgetting rates for sequential input (Bauer et al., 2000). Therefore they might display stronger abstraction due to forgetting (discussed as “fuzzy” representations in Forest et al., 2023). Additionally, young children’s learning (until approx. 6 years) in general seems to benefit from resampling the very same information repeatedly (Horst, 2013; Hudson Kam & Chang, 2009; Pelz & Kidd, 2020). This phenomenon can be related to an early matured (slow) cortical learning system that represents experiences based on repeated encounters, as opposed to a fast learning system (mediated by the hippocampus), which develops later and is thus less available to young learners (Gómez, 2017; Jabès & Nelson, 2015; McClelland et al., 1995; O'Reilly et al., 2014). Thus, spreading out several learning opportunities with the same input over an extended time period, as we did in our study, could favor subsequent generalization to new input also in young children. Additionally, our study design might have aided the formation of explicit knowledge about sequence rules, especially in children: children as young as 6 years of age (5-year-olds in the end of Year 2) reported the same level of rule knowledge as older children and adults (opposed to more explicit rule knowledge in adults vs. children demonstrated by Ferman & Karni, 2010; Hickey et al., 2019; Jung et al., 2020; Smalle, Page, et al., 2017). Wilhelm et al. (2013) showed that children (8-11 years old) benefitted more than adults from one night of sleep in building verbalizable knowledge from their implicit sequence learning experience. Becoming aware of underlying sequence rules, in turn, has been proposed to drive the consolidation of sequence knowledge in an offline period with sleep (Janacsek & Nemeth, 2012).

We show that in a longitudinal setting that allows for rule acquisition across a total of six sessions with the same material before testing transfer, children as young as 5 years at first learning successfully applied the acquired visual rules to completely new surface features (i.e., an unfamiliar stimulus category). This finding extends generalization effects previously reported in development, since prior single-session (Jung et al., 2020; Nowak & Baggio, 2017) and multi-session (Ferman & Karni, 2014) studies on sequence learning in children so far have only shown successful transfer to new instances of the same underlying category (e.g., unseen pictures of the same animals tested after triplet exposure to specific animal pictures in Jung et al., 2020).

Age-dependent representations of sequential regularities and how they might support generalization effects, are discussed as part of Chapter V in the context of broader theories on rule transfer. In this context I will furthermore consider timescales of rule generalization, from sessions spanning one week vs. one year (see comparable transfer effects in 6-year-olds Year 2 [after 6 sessions with stimulus set 1] with the control group of 7-year-olds in Year 1 [after 3 sessions with sessions with stimulus set 1]). A combined review seems useful, because Project 1 also speaks to transfer effects before and after a long-term delay in 7-year-olds vs. adults.

4.5. Conclusion

To conclude, the present study characterized long-term trajectories in visual AGL across different age groups (5-year-olds, 6-year-olds & adults), in two sets of sessions spanning a 12-month delay. Whilst our findings did not confirm that an earlier developmental timing of several AGL instances results in better learning outcomes in the long run, they show that: (1) Children from 6 years onwards successfully learn complex visual sequence rules across several sessions. (2) They use their acquired rule knowledge after a 12-month delay for quicker and additive relearning of the same input compared to before the delay and (3) for transfer to new but related input, both in an adult-like fashion.

Furthermore, the present study has been one of the first attempts to directly test maturational effects against long-term effects of prior exposure to environmental regularities in children. In doing so, our findings substantiate that prior learning experiences can lead to a quicker re-acquisition of sequence rules in relearning after one year. This seemed to be the case even for young children with first task exposure at age 5 years, who had lacked to show any behavioral learning effects before the long-term delay. The present findings have been discussed with regard to developmental changes in learning mechanisms, the concept of "savings" from previous learning experiences, and possible factors driving generalization of rule knowledge in young children. Broader implications from these findings, informing neurocognitive learning models, underlying plasticity processes and sensitive periods in development are discussed in Chapter V. How exactly prior learning is represented on a neural level and what is necessary for its effects to persist, warrants further investigation. Advances in this research field will help to illuminate what mechanisms allow children to efficiently acquire skills over a longer developmental time.

Chapter IV: Associations of repeated statistical learning with memory & language skills (Project 3)

1. Introduction

Exploiting reoccurring regularities in the environment has been put forward as an underlying mechanism in several learning domains, among them most prominently for language acquisition (Erickson & Thiessen, 2015; Romberg & Saffran, 2010). This mechanism has been investigated as statistical learning or implicit learning and is referred to as sequence learning throughout this dissertation (see Chapter I). Early studies (Gómez, 2002; Saffran et al., 1996) mainly focused on the role of sequence learning mechanisms for segmenting continuous speech input into words (reviewed in Aslin, 2017). By now, other language outcomes like grammar/syntax proficiency and reading skills have been related to sequence learning abilities as well (Arciuli & Simpson, 2012; Evans et al., 2009; Kidd, 2012). This line of research has corroborated a link between sequence learning tasks in different modalities and several measures of natural language processing for typically developed populations – both with regard to performance correlations (Conway et al., 2007; Misyak et al., 2010; Misyak & Christiansen, 2012; Smith et al., 2015) and overlapping neural underpinnings (Conway & Pisoni, 2008; Goranskaya et al., 2016; Skosnik et al., 2002). Taking this idea one step further, stimulation studies have provided evidence that classical "language networks" in fronto-parietal cortex, comprising e.g., Broca's region (BA 44/45), are causally involved in sequence learning performance (Uddén et al., 2008; Uddén et al., 2017; Vries et al., 2010). Thus, if sequence learning and language functions recruit the same neural infrastructure for successful behavior, this suggests shared neurocognitive mechanisms and intertwined trajectories across development.

While most of the above studies investigated adults, sequence learning performance in children has been linked to language outcomes as well, especially to emerging literacy: Greater phonological awareness in first and second language acquisition, a central component of emergent literacy skills, was shown to be associated with better performance in visual sequence learning tasks (Bogaerts et al., 2016; Pavlidou & Bogaerts, 2019; Zinszer et al., 2020). Importantly, this relationship in childhood was argued to extend beyond cross-sectional assessments. Rather, an impaired ability to extract sequential information, based on deficient processing in the underlying procedural memory system, was suggested to cause language disorders, such as developmental dyslexia („procedural deficit hypotheses"; Nicolson & Fawcett, 2007, 2011; Ullman, 2004; Ullman & Pierpont, 2005; see, however, West et al., 2021 for a meta-analysis that questions the causal role of a general deficit in procedural memory for language disorders). A meta-analysis quantified the strength of this link by calculating an average weighted effect size of .46 for impaired learning in visual AGL tasks in dyslexic samples, as compared to healthy controls (van Witteloostuijn et al., 2017, including children and adults).

Given all of the evidence above, it seems plausible to assume that sequence learning is connected to language outcomes in the long run (but: see West et al., 2018 for a critical discussion on evidence generated by the „procedural deficit hypotheses"). Bogaerts et al. (2016) provided first long-term evidence in children at risk for developing reading difficulties and controls, by tracking them in two assessments from first to second grade. They showed that sequence learning and reading skills were positively associated across one year in this early phase of reading acquisition, with sequence learning performance explaining significant variance in reading proficiency even when controlling for individual differences in phonological awareness. However, longitudinal assessments looking into this relationship in unselected populations are scarce (Arciuli & Torkildsen, 2012). Including typically developed samples of both, children and adults, in a multi-session study of sequence learning, Smalle, Page, et al. (2017) assessed vocabulary knowledge in a first session of sequence learning. They reported that vocabulary knowledge was positively associated with (linguistic) sequence learning performance in the same session, and across delays of 4 hours, 1 week, and 12 months in a group of 8-9-year-old children (r = .30-.40), but not in adults. Ferman and Karni (2010) demonstrated that phonological aspects of a sequence learning task with syllables are acquired by both children age 8-12 years and adults throughout 15 consecutive sessions and retained after a 2-month retention interval. In their study, adults (and 12-year-old children) were better than 8-year-old children in acquiring these phonological features, providing evidence for this link in adults as well. Their findings furthermore suggest that prior language knowledge might facilitate (linguistic) sequence learning across several sessions. Note that the latter two studies (Ferman & Karni, 2010; Smalle, Page, et al., 2017) reversed the direction of the link put forward by clinical research on dyslexia, and found that (earlier) language skills contributed to (later) sequence learning outcomes.

Closely related to the language domain (Ullman, 2004), memory processes have been intrinsically implicated in sequence learning (Perruchet, 2019; Pothos, 2007; Thiessen, 2017). Thiessen (2017) even put forward that sequence learning can be modelled as a phenomenon which is entirely based on memory processes like chunking (reviewed also in Perruchet 2019), (re)activating, integrating (consistent) and forgetting (interfering) pieces of information. He stated that this "extraction and integration" approach to sequence learning means that learning “does not involve explicit computation of statistics, but rather the extraction of elements of the input into memory traces, and subsequent integration across those memory traces that emphasize consistent information” (Thiessen, 2017, p. 1). From equating sequence learning mechanisms with memory processes, declarative memory encoding and retrieval skills can be hypothesized to contribute to sequence learning outcomes to varying degrees, depending on age.

An early matured procedural memory system as opposed to a continuously maturing declarative memory system (Finn et al., 2016; Meulemans et al., 1998; Parkin & Streete, 1988) seems to be mainly recruited for learning in childhood (reviewed in Gualtieri & Finn, 2022). Due to these developmental constraints, declarative memory skills might be less relevant for successful sequence learning in younger age (for a comprehensive discussion on memory development and sequence learning, see Forest et al., 2023). In contrast, adults have been shown to rely predominantly on their declarative memory system in learning situations (Finn et al., 2014; Ullman, 2001) – a reliance that has been successfully manipulated by cognitive load and brain stimulation in the context of sequence learning tasks (Ambrus et al., 2020; Smalle et al., 2022). Based on the above literature, one would predict a stronger association between adults’ declarative memory skills and sequence learning performance as compared to children.

On a shorter timescale, working memory has been discussed to at least “modulate or gate” (Conway, 2020, p. 279) the learning of sequential regularities, as reviewed recently across different task paradigms. Putting the involvement of memory functions to the test, studies have manipulated different aspects of working memory and tested consequences for sequence learning outcomes. One study of interest investigated several working memory characteristics in three experiments with adults, using a speech segmentation paradigm (Palmer & Mattys, 2016): They report that a slower presentation rate improved task performance and found comparable effects of a visual and an auditory two-back task during exposure, which both impeded processing of the auditorily presented speech regularities. Hendricks et al. (2013) tested working memory involvement in visual sequence learning, using a similar dual-task setting. They aimed to dissociate influences on sequence learning from working memory load during exposure vs. during test in adults. In their study, encoding of visual regularities was observed even under dual task conditions, while the same conditions during test lead to a break down in discrimination performance. However, in a variation of their experimental set-up that used new (transfer) stimuli in the test phase, performing another task during encoding impaired test performance in their visual sequence task as well. These two studies stress that working memory capacities seem to be involved in sequence learning (1) in a domain-general manner at (2) different processing stages, and (3) might particularly matter for generalizing encountered regularities to new input.

Regarding possible developmental differences for this link, a cross-sectional study on visual sequence learning in children and adults reported that working memory skills were associated with task performance only in adults (r = .30-.40), but not in children age 7-9 years (Arnon, 2019). The only study to our knowledge that investigated memory involvement in a multi-session setting showed similar age differences (Smalle, Page, et al., 2017): Working memory capacity assessed at the beginning of the study was associated with sequence learning outcomes across up to 12 months only in adults (r = .50-.60), not in children age 8-9 years.

All of the above studies are based on viewing sequence learning abilities as a separate, well-defined characteristic of individuals that can be reliably measured and related to interindividual differences in language and memory outcomes in a meaningful fashion. In support of this notion, extracting sequential regularities has been described as a trait-like ability in its unique relationships to various cognitive, academic and personality measures in a large sample of adolescents (Kaufman et al., 2010). In the same vein, differences in sequence learning performance have been found to explain a unique proportion of behavioral variance in language outcomes of children (Bogaerts et al., 2016; Pavlidou & Bogaerts, 2019; Zinszer et al., 2020) and adults (Conway et al., 2010; Danner et al., 2017; Kaufman et al., 2010; Misyak & Christiansen, 2012), even after accounting for differences in other cognitive domains like general intelligence, processing speed, or working memory.

However, low reliabilities of sequence learning measures have been repeatedly criticized (Arnon, 2019; Frost et al., 2019; Siegelman et al., 2017; West et al., 2018), which lead the respective authors to question how useful it is to relate performance in these tasks to interindividual differences in other cognitive domains. Studies have estimated the reliability of sequence learning tasks on different levels, both in developmental and adult samples: A lack of statistically significant correlations across equivalent tasks in different modalities (Arnon, 2019; Pavlidou & Bogaerts, 2019) suggests that sequence learning performance is not associated within individuals across the visual, auditory and tactile domain – at least not in children. In the following, we will focus on accuracy-based reliability measures in visual sequence tasks, as they are of interest for the current study. Split-half or internal reliabilities of these tasks were reported to range from close to zero (r = .04 Zinszer et al., 2020 in children) to small/moderate (adults: r = .40 in Danner et al., 2017; Kaufman et al., 2010, r = .69-.75 in Farkas et al., 2023; children: r = .40 in Bogaerts et al., 2016; r = .50 in West et al., 2018 [see non-verbal Hebb task]; r = .46-.72 in Arnon, 2019) to reasonably high strengths (adults: r = .72-.91 in Arnon, 2019:, r = .83-.88 in Siegelman et al., 2017; children: $r \geq .80$ in Torkildsen et al., 2019; Qi et al., 2019). Available estimates of test-retest reliabilities focus on adults and provide evidence for moderate performance correlations of r = .45 (Arnon, 2019) to r = .68 (Siegelman et al., 2017) in visual sequence tasks across two to three months (but: same test-retest correlation of .01 [n.s.] for 7-9-year-olds in Arnon, 2019). Extending the test-retest interval to 12 months, a first estimate (r = .52) across this long delay comes from adult reaction time data in a visuomotor sequence task (see discussion in (Farkas et al., 2023); but see also their critical view on the role of consolidation in test-retest measurements).

As illustrated, reliability estimates for sequence learning tasks vary widely. Thus, a first aim of this exploratory approach in Project 3 was to check the reliability of response accuracy in a multi-session AGL task across one year in our combined sample of different age groups (5-year-olds, 6-year-olds, 7-year-olds, Adults 1, Adults 2, see Table 1 for study design of all groups). To this end, associations between AGL task performance before and after a 12-month delay were calculated as a proxy for test-retest reliability of the applied AGL task in the study setting of this dissertation (task design for groups of Project 1 & Project 2, see Fig. 1 & Fig. 8). Additionally, AGL task performance within each year was correlated between the two stimulus sets used, to provide an estimate for parallel test reliability in two successive sessions. Both stimulus sets implemented the same underlying rule set in an equivalent AGL task, but displayed it with different visual surface features (pictures of different stimulus categories).

The main goal of Project 3 then was to test associations of sequence learning (AGL task performance) with memory and language skills of children and adults on different multi-session timescales in an exploratory manner: This included relationships across one week (both assessed within Year 1), across a delay of 12 months (predictive value of cognitive skills from Year 1 to AGL in Year 2) and for relearning after the delay (both assessed within Year 2).

Based on the literature reviewed above, we expected better working memory and declarative memory retrieval to facilitate performance in our sequence learning task across one week and after the one-year delay. Since the sequence task was tailored to the working memory capacities of children 5-8-years, respective associations might emerge in our child groups as well, opposed to a previous multi-session study that used stimulus material matched to adult memory capacities (Smalle, Page, et al., 2017). For declarative memory skills (encoding and retrieval after a short delay), adults were expected to show stronger positive associations with AGL performance than children. This hypothesis is based on reports in adults relying mainly on the declarative memory system in learning situations. Language grammar skills were hypothesized to be positively related to AGL outcomes in all age groups, given previous findings on positive associations between language processing and sequence learning performance in children and adults.

All hypotheses were tested in an exploratory manner, keeping in mind the main focus of this dissertation was to compare groups with a different developmental timing of several AGL instances in their learning outcomes (see Project 1 & Project 2), using a cohort study approach. Our study design was consequently optimized for this purpose, resulting in limited power to look into interindividual differences in learning and how they relate to other cognitive domains. Nevertheless, this chapter aims to provide a first multi-session perspective on the link between language and memory skills to visual AGL in children and adults.

2. Methods

2.1. Participants

In Project 3, we combined data of all available age groups of Project 1 and Project 2 to look into associations between visual sequence learning performance and cognitive skills. This involved 5-year-olds, 6-year-olds, 7-year-olds (5/6/7 years old ± 2 months at Session 1) and two adult groups (Adults 1, Adults 2).

Depending on the data required for each set of analyses (correlations within Year 1 vs. across one year vs. within Year 2, see *Results* sections), all age groups with available data we considered. Due to slight differences in the study design between these age groups (see Table 1), this meant that for some analyses only subgroups of the whole sample could be considered. The objective was to optimize statistical power for all associations tested. Within this rationale, all datasets with at least 3 out of 5 AGL task blocks per session (inclusion criterion for all AGL analyses of the dissertation, see Chapter II) and available psychometric scores in the cognitive variables at the timepoint of interest (Session 1 in Year 1 or Session 4 in Year 2, see section *Assessment of memory and language skills*) were included. The resulting descriptive information (AGL parameters & cognitive scores) are detailed for Year 1 and Year 2 separately (see Table 2-5), for all respective age groups (see Chapter II & III for additional sample characteristics like exact age and session timing). Additionally, sample sizes are indicated in the results tables below, which display the correlations per results section.

The groups of 7-year-olds and Adults 1 were comprehensively described in Chapter II. For the analyses of the present chapter, we included data of 27 seven-year-olds and 28 adults for analyses relating sequence learning performance to cognitive assessments in Year 1 (descriptive data see Table 18 & 19). For predicting AGL performance in Year 2 from cognitive scores in Year 1, we included all available AGL data from the home follow-up from these groups (descriptive data see Table 20 & 21), entailing 16 seven-year-olds and 20 Adults 1. 7-year-olds and Adults 1 considered for Year 1 and Year 2 here, correspond to the same participants for whom learning trajectories were discussed before in Project 1 (see Table 2 & Table 3 of Chapter II for additional sample characteristics like exact age and session timing).

For the three additional groups investigated in Project 2 (comprehensively described in Methods of Chapter III), here we included data of 27 five-year-olds, 28 six-year-olds and 29 Adults 2 for analyses relating sequence learning performance to cognitive assessments within Year 1 (descriptive data see Table 18 & 19). For predicting AGL performance from cognitive scores in Year 1 and analyses within in Year 2, relating sequence (re)learning performance after the one-year delay to cognitive assessments, we analyzed available data of 22 five-year-olds, 25 six-year-olds and 18 adults. These samples correspond to all participants for whom learning trajectories were discussed in Project 2, except for 2 participants per age group, who completed all sessions of Year 2 at home due to ongoing COVID-19 health restrictions (*n* = 2 five-year-olds, 6-year-olds, Adults 2, respectively). For those participants, we were not able to collect cognitive assessments in the lab, leading to their exclusion for correlational analyses with AGL performance. For single cognitive variables and timepoints, not all children completed the respective subtests and thus were excluded for the respective correlation (*n* = 1 six-year-old in Year 1 for Declarative Memory and German Grammar I, as indicated in the footnotes of Table 2; *n* = 2 five-year-olds & *n* = 2 six-year-olds in Year 2 for Working Memory, see footnotes respective results tables).

All procedures for participant compensation and ethics approval were described in Chapter II.

2.2. Material and procedure

2.2.1. Multi-session visual sequence learning

Details of the modified AGL task and study design used to measure multi-session visual sequence learning are provided in Chapter II. Here, we shortly summarize the relevant information for Project 3.

Seven-year-olds and Adults 1 completed a total of four sessions (Session 1,2, 3, Transfer 1) in Year 1 with the visual sequence learning task on separate days over the time of approx. one week, and an equivalent set of four additional sessions at home in Year 2 (Session 4, 5, 6, Transfer 2) after a one-year delay (see study design Table 1). For Project 3, we were mainly interested in how AGL parameters (Session 1 to 3, Transfer 1) relate to cognitive scores in psychometric measures for language and memory, both assessed within Year 1 (see *Data analysis*). Additionally, for these groups, equivalent AGL performance from the home follow-up in Year 2 (Session 4 to 6, Transfer 2) was related to previously assessed cognitive skills in Year 1.

Five-year-olds, 6-year-olds and Adults 2 completed a total of three sessions with the visual sequence learning task in Year 1 (Session 1 to 3), and an equivalent of three sessions (Session 4 to 6) with a subsequent transfer session (Transfer 2) in Year 2, each set spread out over the time of approx. one week (see study design Table 1). Here, we were mainly interested in how AGL parameters in Year 1 (Session 1 to 3) and Year 2 (Session 4 to 6, Transfer 2) relate to cognitive scores in psychometric measures for language and memory, as assessed within the same year (in Session 1 for Year 1 & Session 4 for Year 2, respectively). In addition to this within-year focus, all performance differences from Year 2 (concerning Session 4 to 6, Transfer 2) were related to previously assessed cognitive skills in Year 1 (see *Data analysis*).

The AGL task used for measuring visual sequence learning on a tablet computer is described in detail in Chapter II. It entailed 5 task blocks per session, consisting of alternating phases of learning (exposure to 18 grammatical sequences each) and test (10 trials of two-alternative forced choice responses between one grammatical and one ungrammatical sequence). This amounted to an AGL task exposure of approx. 25-30 minutes per session. The first 3 sessions of each year used a first stimulus set (stimulus set 1), while the last session of each year (7-year-olds & Adults 1) or Year 2 (5-year-olds, 6-year-olds and Adults 2), respectively, employed a second stimulus set (stimulus set 2) to investigate transfer of learned AG rules.

2.2.2. Assessment of memory and language skills

To assess declarative memory, working memory and German grammar skills, we administered equivalent psychometric tests in Session 1 (Year 1) for two groups (7-year-olds & Adults 1), and in Session 1 (Year 1) and Session 4 (Year 2) for three groups (5-year-olds, 6-year-olds, Adults 2), see Table 18-21. 7-year-olds and Adults 1 completed the AGL task in Year 2 as a home follow-up, where no cognitive assessments could be administered.

2.2.2.1. Declarative memory

Memory skills of encoding and retrieving new information were measured using subtests of the “Kaufman Assessment Battery for Children II” (KABC II, subtests “Atlantis” (encoding) and “Atlantis – Abruf nach Intervall” (retrieval); adapted German version by Melchers & Melchers, 2015) for children and subtests of the “Kaufman-Test zur Intelligenzmessung für Jugendliche und Erwachsene“ (K-TIM; subtests “Symbole lernen” (encoding) and “Symbole – Abruf nach Intervall”; adapted German version by Melchers, Schürmann, & Scholten, 2006) for adults. Participants of all ages had to learn associations between words and drawings that had to be retrieved in a surprise cued recall test after a delay of 15 to 20 minutes (children) or 20 to 30 minutes (adults), respectively. For adults, the task included learning and recalling language-like rules embedded in these associations, like using the past tense of verbs for added features in the drawings. Split-half reliabilities of the administered subtests were .91-.97 for children and .97 for adults, as indicated in the psychometric manuals.

2.2.2.2. Working memory

For assessing working memory capacity, children and adults completed digit span tasks, which meant repeating verbally presented digit spans of increasing length. In children, this was done by administering the subtest “Zahlen nachsprechen“ from the KABC II (Melchers & Melchers, 2015). For adults, the administered working memory subscale of the “Wechsler Adult Intelligence Scale“ (WAIS-IV; adapted German version by Petermann, 2012) included additional tasks, which asked them to repeat the presented digit spans in reverse order or reorder all heard digits according to their chronological order. The administered measures were reported to have a split-half reliability .81 for children and internal consistency (Cronbach’s α) of .93 for adults, as detailed in the psychometric manuals.

2.2.2.3. German grammar

Two aspects of German grammar skills were assessed in children and adults, involving tasks that tested plural forms for various nouns (German Grammar I) and tasks that tested more general grammar rules concerning e.g., syntax and tenses (German Grammar II). In children, these tasks were verbally presented and taken from the “Sprachstandserhebungstest für Kinder im Alter von 5 und 10 Jahren” (Subscale 8 and 9 of the SET 5-10; Petermann, 2018). Their tasks testing general grammar rules increased in complexity from 6 to 7 years of age (up to 6 years: judging presented sentences as grammatically correct or incorrect; from 7 years onwards: correcting incorrect sentences). Adults completed the subtest “Pluralbildung” and the subscale “Grammatik” from the “Deutschtest für die Personalauswahl” (D-PA; Rieser & Liepmann, 2014) in a written form that included time limits for each task block. Reliability coefficients of the administered subtests, taken from the psychometric manuals, were 94-.96 for adults (Cronbach’s α for internal consistency & split-half reliabilities) and .71-.84 for children (Cronbach’s α).

All test scores for children and adults were normalized according to age (and additionally according to educational level, operationalized as highest school degree for German Grammar II in adults). This was done except for German Grammar I in adults, for which norms were not available and for which hence raw scores were analyzed. In addition to normalized test scores, we considered the maximum number of digits recalled in the working memory task as an additional measure for working memory capacity across all age groups.

Table 18 (child groups) and Table 19 (adult groups) show the administered subscales of these psychometric tests together with the achieved test scores and sequence learning parameters (session differences in the visual AGL task) for all age groups in the final sample of Year 1. Table 20 and 21 list the same information for Year 2, entailing sequence learning parameters (AGL task) for all age groups and cognitive subtests with achieved scores. Cognitive assessments in Year 2 were available for 5-year-olds, 6-year-olds and Adults 2 only, as described above.

Table 18

Year 1: Assessed Memory & Grammar Skills, AGL Parameters with Means (SD) in Children

Cognitive Skill / AGL Parameter		**5-Year-Olds (*n* = 27)**	**6-Year-Olds (*n* = 29)[3]**	**7-Year-Olds (*n* = 27)**
Declarative Memory	Atlantis (Encoding)[1]	11.07 (3.01)	11.75 (3.11)	11.67 (2.76)
	Atlantis Delayed (Retrieval)[1]	12.48 (2.29)	11.96 (2.41)	12.37 (2.39)
Working Memory	Number Recall[1]	10.93 (2.84)	10.21 (2.73)	10.48 (2.62)
	Max. Number of Digits recalled	4.11 (.80)	4.38 (.73)	4.67 (.68)
German Grammar I (Plural)	Subscale 8 SET[2]	53.30 (8.54)	62.00 (9.90)	57.56 (9.66)
German Grammar II (General)	Subscale 9 SET[2]	56.67 (12.91)	69.50 (12.15)	60.19 (15.90)
Learning Gains AGL	Session 3 – Session 1	.03 (.14)	13 (.11)	.10 (17)
Transfer Savings AGL	Transfer 1 – Session 1	NA	NA	.08 (.15)
Transfer Loss AGL	Transfer 1 – Session 3	NA	NA	-.02 (.15)

Note. [1] KABC-II (Melchers & Melchers, 2015): normalized to *M*±*SD* of 10±3; [2] SET 5-10 (Petermann, 2018): normalized to *M*±*SD* of 50±10; [2] SET 5-10 (Petermann, 2018): normalized to *M*±*SD* of 50±10; [3] *n* = 1 missing value for subtests Atlantis, Atlantis Delayed, Subscale 9 SET; AGL: difference in proportion correct of the Artificial Grammar Learning Task; NA: not applicable due to study design.

Table 19

Year 1: Assessed Memory & Grammar Skills, AGL Parameters with Means (SD) in Adults

Cognitive Skill / AGL Parameter		**Adults 1 (*n* = 28)**	**Adults 2 (*n* = 29)**
Declarative Memory	Rebus (Encoding)[1]	11.71 (2.17)	12.07 (2.10)
	Rebus Delayed (Retrieval) [1]	11.86 (1.53)	11.48 (1.82)
Working Memory	Digit Span[2]	11.11 (2.25)	10.38 (2.54)
	Max. Number of Digits recalled	6.79 (.96)	6.62 (.90)
German Grammar I (Plural)	Subtest AR4 D-PA[3]	12.25 (1.11)	12.28 (.96)
German Grammar II (General)	Subscale AG D-PA[3]	53.46 (9.16)[a]	58.62 (25.30)[a]
Learning Gains AGL	Session 3 – Session 1	.13 (.10)	.12 (10)
Transfer Savings AGL	Transfer 1 – Session 1	.10 (.13)	NA
Transfer Loss AGL	Transfer 1 – Session 3	-.04 (.07)	NA

Note. [1] K-TIM (Melchers et al., 2006): normalized to *M*±*SD* of 10±3; [2] WAIS-IV (Petermann, 2012): normalized to *M*±*SD* of 10±3*;* [3] D-PA (Rieser & Liepmann, 2014): Subscale AG normalized to *M*±*SD* of 50±10, Subtest AR4: no age norms available (raw scores); AGL: difference in proportion correct of the Artificial Grammar Learning Task; NA: not applicable due to study design.

[a] n = 3 adults with $T < 15$ (n = 1 Adults 1, n = 2 Adults 2). Result patterns for correlations were checked for excluding these participants (see respective *Results* tables).

Table 20

Year 2: Assessed Memory & Grammar Skills, AGL Parameters with Means (SD) in Children

Cognitive Skill / AGL Parameter		**5-Year-Olds (n = 22)[3]**	**6-Year-Olds (n = 25)[3]**	**7-Year-Olds (n = 16)**
Declarative Memory	Atlantis (Encoding)[1]	11.54 (2.47)	11.74 (3.10)	NA
	Atlantis Delayed (Retrieval)[1]	12.58 (2.47)	11.96 (2.39)	
Working Memory	Number Recall[1]	11.08 (2.93)	10.37 (2.56)	NA
	Max. Number of Digits recalled	4.65 (.67)	5.09 (.60)	NA
German Grammar I (Plural)	Subscale 8 SET[2]	53.92 (8.71)	61.67 (9.27)	NA
German Grammar II (General)	Subscale 9 SET[2]	56.75 (12.11)	68.41 (12.15)	NA
Retention Year 1 to Year2 AGL	Session 4 – Session 3	.00 (.13)	-.03 (.13)	.02 (.13)
Learning Gains AGL	Session 6 – Session 4	.12 (.13)	.11 (.11)	.09 (.11)
Transfer Savings AGL	Transfer 2 – Session 4	.06 (.11)	.03 (.12)	.02 (.10)
Transfer Loss AGL	Transfer 2 – Session 6	-.06 (.12)	-.08 (.12)	-.06 (.11)

Note. [1] KABC-II (Melchers & Melchers, 2015): normalized to $M \pm SD$ of 10±3; [2] SET 5-10 (Petermann, 2018): normalized to $M \pm SD$ of 50±10; [3] n = 2 missing values for subtest Number Recall per age group; AGL: difference in proportion correct of the Artificial Grammar Learning Task; NA: no assessment due to home follow-up.

Table 21

Year 2: Assessed Memory & Grammar Skills, AGL Parameters with Means (SD) in Adults

Cognitive Skill / AGL Parameter		**Adults 1 (*n* = 20)**	**Adults 2 (*n* = 18)**
Declarative Memory	Rebus (Encoding)[1]	NA	12.30 (2.30)
	Rebus Delayed (Retrieval)[1]		11.60 (1.82)
Working Memory	Digit Span[2]	NA	10.80 (2.59)
	Max. Number of Digits recalled	NA	6.22 (1.00)
German Grammar I (Plural)	Subtest AR4 D-PA[3]	NA	12.20 (.77)
German Grammar II (General)	Subscale AG D-PA[3]	NA	54.61 (7.75)
Retention Year 1 to Year2 AGL	Session 4 – Session 3	-.03 (.06)	-.03 (.07)
Learning Gains AGL	Session 6 – Session 4	.02 (.05)	.05 (.07)
Transfer Savings AGL	Transfer 2 – Session 4	.01(.05)	.01 (.08)
Transfer Loss AGL	Transfer 2 – Session 6	-.02 (.05)	-.05 (.07)

Note. [1] K-TIM (Melchers et al., 2006): normalized to *M*±*SD* of 10±3; [2] WAIS-IV (Petermann, 2012): normalized to *M*±*SD* of 10±3; [3] D-PA (Rieser & Liepmann, 2014): Subscale AG normalized to *M*±*SD* of 50±10, Subtest AR4: no age norms available (raw scores); AGL: difference in proportion correct of the Artificial Grammar Learning Task; NA: no assessment due to home follow-up.

2.3. Data Analysis

As in the previous chapters, we characterized AGL performance scores as proportion of correct test trials: For each session, learning was assessed as the mean performance of the total of 50 test trials completed.

AGL trials with reaction times shorter than 200 ms were disregarded, since we did not consider it feasible to successfully process the two sequences within less than this time. This exclusion criterion reduced trial numbers for AGL sessions of all age groups in Year 1 by 0.20 % (a total of 54 excluded trials in 5-year-olds, 6-year-olds & 7-year-olds, with a maximum of 16 trials excluded per subject) and for AGL sessions of all age groups in Year 2 by 0.26 % (a total of 64 excluded trials in 5-year-olds, 6-year-olds, 7-year-olds & Adults 1, with a maximum of 12 trials excluded per subject).

The following AGL session differences were considered as measures for task performance to be correlated with cognitive scores, corresponding to the AGL parameters reported in Chapter II and Chapter III:

1. For associations within Year 1, we calculated the following difference scores to quantify improvement with stimulus material 1 and transfer across one week to stimulus material 2 in the AGL task:
    - Session 3 – Session 1 (Learning Gains)
    - Transfer 1 – Session 1 (Transfer Savings)
    - Transfer 1 – Session 3 (Transfer Loss)
2. For associations including AGL in Year 2, we calculated equivalent differences to quantify improvement with stimulus material 1 and transfer across one week to stimulus material 2 in the AGL task. In addition, consolidation in stimulus set 1 from the last session in Year 1 to the first session in Year 2 was analyzed:
    - Session 4 – Session 3 (Retention Year 1 to Year 2)
    - Session 6 – Session 4 (Learning Gains)
    - Transfer 2 – Session 4 (Transfer Savings)
    - Transfer 2 – Session 6 (Transfer Loss)

Pearson correlations were calculated and their coefficients are reported as significant for two tailed *p*-values < .05, corrected for multiple comparisons by controlling the false discovery rate (FDR, Benjamini & Hochberg, 1995). If scores were not normally distributed as assessed by the Shapiro-Wilk Test for bivariate normality, Spearman correlation coefficients (r_s) were calculated and are reported with FDR-corrected *p*-values. Multiple comparison corrections were applied per reported age group (within children's group & adult group, detailed in all results tables), within each timescale (separately for associations within Year 1, across one year & within Year 2) and there within each AGL performance difference (i.e., for all 5 associations of cognitive skills with e.g., Learning Gains AGL in children in Year 1, see Table 7 left column). Since this exploratory approach in Project 3 was based on rather small sample sizes with limited power for individual differences analyses, we marked correlations with an absolute size of $| r | \geq .30$ independent of their statistical significance as trends. Additionally, we report the Bayes factor (*BF*, see explanation below) for each calculated correlation (see results tables for *BFs* > 3 & Appendix D for all *BFs* per association)**.**

Data analyses were performed in the software R (Version 4.1.0; R Core Team, 2021) and JASP (Version 0.14.1; JASP Team, 2021), the latter using default priors and reporting the BF_{10}. The *BF* helps evaluating whether the data at hand support the null-hypothesis (H_0, in this case that the correlation is zero) or the alternative hypothesis (H_1). This approach has been used in studies with a similar correlational approach with small sample sizes before (Pavlidou & Bogaerts, 2019). *BF* values between 1/3 and 1/10 indicate moderate evidence for the H_0, while a BF of lower than 1/10 is considered strong evidence for the H_0; a BF between 1 and 1/3 is defined as anecdotal evidence for the H_0 (Schönbrodt & Wagenmakers, 2018). *BF* values between 3 and 10 indicate moderate evidence for the H_1, while a *BF* from 10 onwards is considered as strong evidence for the H_1 and a *BF* between 1 and 3 is defined as anecdotal evidence for the H_1 (Schönbrodt & Wagenmakers, 2018).

3. Results

3.1. Estimating test-retest and parallel test reliabilities in the present sequence learning task

We estimated two measures for the present visual AGL task across one year to evaluate test-retest reliability: Correlations were calculated between session performance levels in the first stimulus set (Session 1 ~ Session 4) and the second stimulus set (Transfer 1 ~ Transfer 2; see Table 22) across all age groups with available data for these sessions.

Table 22

AGL Estimations for Test-Retest and Parallel Test Reliability per Available Age Group

AGL Reliability Measure	**Whole sample (n = 103)**	**7-year-olds & Adults 1 (n indicated below)**
	Session 4 ~ Session 1 (Stimulus Set 1)	Transfer 1 ~ Transfer 2 (Stimulus Set 2, n = 36)
Test-Retest Reliability (Year 1 ~ Year 2)	.77** [.60-.84]	.79** [.62-.89]
	Transfer 2 ~ Session 6 (Year 2)	Transfer 1 ~ Session 3 (Year 1, n = 55)
Parallel Test Reliability (Stimulus Set 1 ~ Stimulus Set 2)	.84** [.77-.89]	.84** [.74-.90] [9]

Note. [...] = 95% Confidence Interval (CI), **= p <.001 & BF_{10} > 100.

Averaged performance in Session 1 (Year 1) was positively correlated with performance in Session 4 (Year 2) for AGL with stimulus set 1 (r = .77, 95% CI = .60-.84, n = 103, p < .001, BF_{10} > 100, see Fig. 17). Additionally, we checked whether test-retest performance correlations with stimulus set 1 would hold for the last session in Year 1 (Session 3) and the first session in Year 2 (Session 4): This correlation was of similar magnitude (r = .84, 95% CI =.77-.89, n = 103, p < .001, BF_{10} > 100).

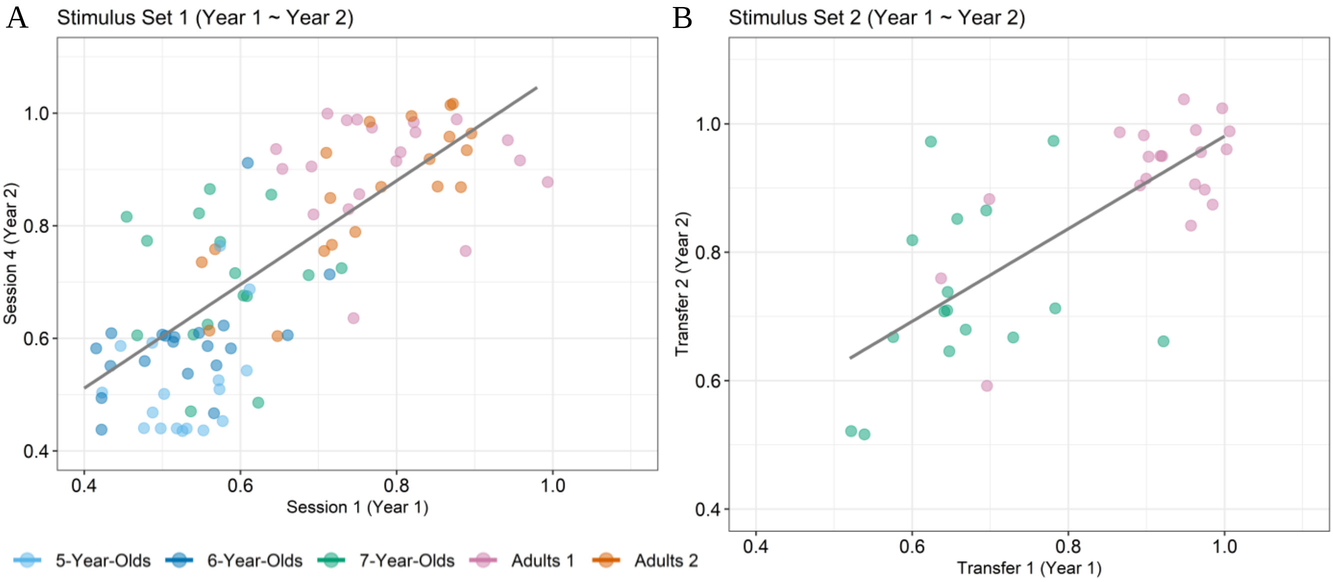
In the subsample of 7-year-olds and Adults 1 with available data in two (transfer) sessions with stimulus set 2, AGL performance in Transfer 1 (Year 1) correlated with

[9] A correlation within a subset of n = 36, which corresponds to the sample for Test-Retest Reliability in the line above (see Transfer 1 ~ Transfer 2 (Stimulus Set 2)), replicated this correlation (Pearson's r = .77, 95% CI = .59-.87, p < .001, BF_{10} > 100).

performance in Transfer 2 (Year 2) by r = .79 (95% CI = .62-.89, n = 36, p < .001, BF_{10} > 100, see Fig. 17).

Figure 17

Test-Retest Reliability in the Multi-Session AGL Task for Stimulus Set 1 & 2



Note. Pearson correlations (within-subject) between proportion correct in the test phase for Session 1 (Year 1) with Session 4 (Year 2) using stimulus set 1 (A), and for Transfer 1 (Year 1) with Transfer 2 (Year 2) using stimulus set 2, respectively (B). Both stimulus sets used the same grammar, but different picture categories (animals or colors). Dot colors indicate subjects' age (see legend). Gray lines display the linear regression.

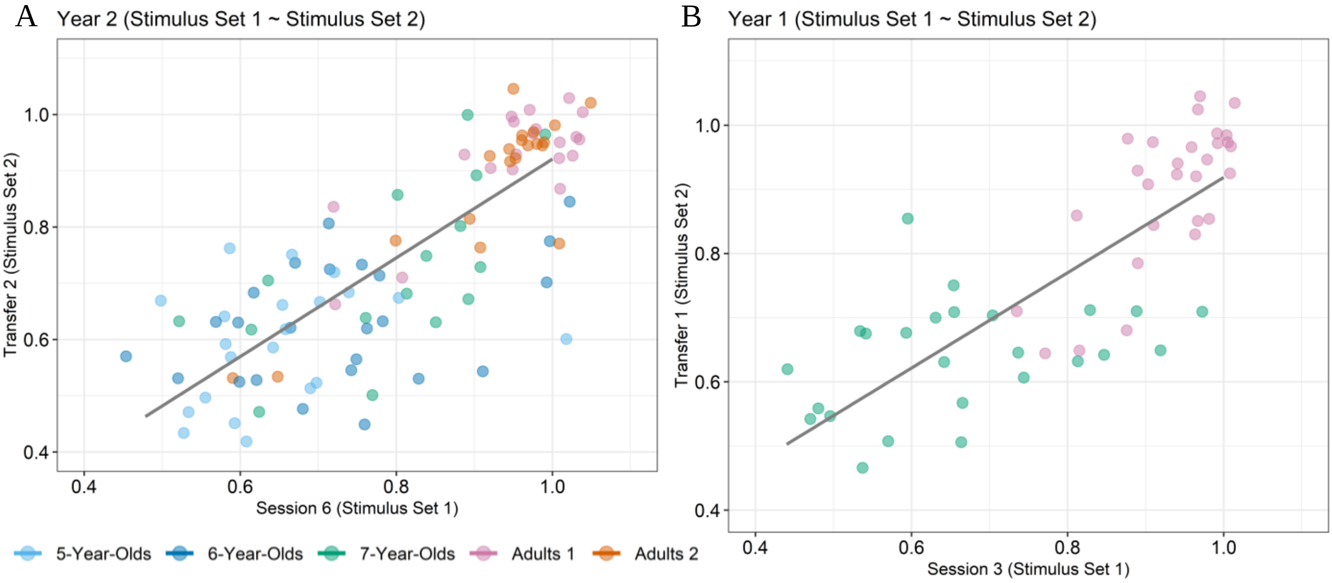
As an additional assessment of task reliability in our setting, we correlated AGL performance correlations between the two stimulus sets, using the two last Sessions of Year 2 (Session 6 and Transfer 2, see Table 22). The resulting correlation of r = .84 (95% CI = .77-.89 , n = 103, p < .001, BF_{10} > 100, see Fig. 18) provides a proxy for the parallel test reliability between stimulus set 1 and stimulus set 2 in two subsequent sessions (separated by 2 days on average), which should be less prone to (long-term) consolidation effects and additional task exposure between test-retest AGL sessions. This proxy for parallel test reliability of our AGL task between stimulus set 1 and stimulus set 2 was replicated in the subsample of 7-year-olds and Adults 1 for the equivalent sessions in Year 1 (Session 3 and Transfer 1: $r_{s\,=}$ .84 (95% CI = 74-.90), p < .001, BF_{10} > 100, see Fig. 18).

To sum up, performance in our AGL task across one year was consistently correlated (r = .77-.84) in a mixed sample of children and adults within the same stimulus sets, providing a first estimate of long-term test-retest reliability in this multi-session setting. The size of these test-retest measures was confirmed by parallel test estimates of the same magnitude (*r* = .77-.84), which emerged for performance across the two stimulus sets with the same underlying rule set, assessed in subsequent sessions with a delay of only several days.

Correlations between Year 1 and Year 2 (Session 4 ~ Session 1) for the repeated assessment of cognitive skills with psychometric measures in our sample ranged from *r* = .17 to *r* = .89 for raw scores of memory and language assessments (children: .17-.55, adults: .27-.89). These values were partly below the reliability coefficients of the test norms reported in the psychometric manuals, especially for children, and are listed in detail in Appendix D. Respective correlations could only be calculated for groups, who completed two psychometric assessments over the course of the study (5-year-olds, 6-year-olds & Adults 2).

Figure 18

Parallel Test Reliability in the Multi-Session AGL Task in Year 1 & 2



Note. Pearson correlation (within-subject) between proportion correct in the test phase for Session 6 (stimulus set 1) with Transfer 2 (stimulus set 2) in Year 2 (A). Spearman correlation for Session 3 (stimulus set 1) with Transfer 1 (stimulus set 2) (B). Both stimulus sets used the same grammar, but different picture categories (animals or colors). Dot colors indicate subjects' age (see legend). Gray lines display the linear regression.

3.2. Are memory and grammar skills in Year 1 associated with multi-session sequence learning performance?

To test for a possible association of memory and language skills assessed in Year 1 with visual sequence learning, we calculated correlations of AGL task performance with these variables on two timescales. This was done (1) within Year 1, relating cognitive scores assessed in the beginning of the study (Session 1) to multi-session AGL performance across one week, and (2) across one year, relating cognitive scores assessed in the beginning of the study (Session 1) to multi-session AGL performance after the one year delay, i.e., in Year 2.

Given the exploratory approach of this chapter, correlations with an absolute size of $|r| \geq .30$ independent of their statistical significance are reported in all *Results* sections (with *BF*s included; see also *Data Analysis*), and will be discussed as trends in the *Discussion*.

3.2.1. Associations within Year 1 for cognitive skills and AGL performance

Within Year 1, we used three AGL task performance scores that quantified performance gains across three sessions with stimulus material 1 in all age groups (Session 3 – Session 1) and transfer effects to stimulus material 2 across one week in 7-year-olds and Adults 1 (Transfer Savings: Transfer 1 – Session 1; Transfer Loss: Transfer 1 – Session 3; for details *Data Analysis*). This was done to keep in line with the Project 1 and Project 2 that used the same session comparisons to look into age differences in AGL across one week (Chapter II & Chapter III).

Table 23 shows all correlations for these three AGL performance parameters with German grammar skills, declarative working memory and working memory scores in Year 1 (see Appendix D for all Scatterplots & *BF*s). A higher working memory capacity was associated with more improvement across three sessions in the first stimulus set across all age groups (longest digit span recalled ~ Session 3 – Session 1; $r_s = .23$, $p = .018$, $BF_{10} = 7.75$).

Numerically, better working memory skills, assessed as scores standardized for age, were also associated with more generalization to the second stimulus set in 7-year-old children (Transfer 1 – Session 1; $r_s = .41$, $BF_{10} = 2.79$). Additionally, 7-year-olds with better German grammar scores (Grammar II, Plural) displayed a more pronounced drop in performance in the transfer session compared to the last session with the first stimulus set compared to 7-year-olds with lower grammar skills (Transfer 1 – Session 3; $r_s = -.32$, $BF_{10} = 1.06$): However, these two associations in 7-year-olds did not remain statistically significant after correcting for multiple comparisons (both corrected $p \geq .175$). Note that for 5-year-olds, 6-year-olds and Adults 2, associations involving AGL transfer in Year 1 could not be evaluated, because these groups were first introduced to the second stimulus set used in the transfer session in Year 2, not in Year 1. For Adults 1, none of the associations between cognitive skills and AGL performance in Year 1 exceeded a value of |.30| or reached statistical significance (all $|r_s| \leq .25$, all corrected $p \geq .452$, all $BF_{10} \leq .68$).

Table 23

Correlations for Memory & Grammar Skills in Year 1 with AGL Performance in Year 1

	Learning Gains AGL (Session 3 – Session 1)		**Transfer Savings AGL** (Transfer 1 – Session 1)		**Transfer Loss AGL** (Transfer 1 – Session 3)	
	Children (*n* = 83)	Adults (*n* = 57)	7-year-olds (*n* = 27)	Adults 1 (*n* = 28)	7-year-olds (*n* = 27)	Adults 1 (*n* = 28)
German Grammar I (Plural)	.18	-.07	.03	-.19	***-.32***	.07
German Grammar II (General)	.21 [a]	.10 [b]	.02 [b]	-.14 [b]	-.06 [b]	-.22
Declarative Memory						
Encoding	-.02 [a]	-.06	.12	-.23	.22	-.25
Retrieval	-.01 [a]	.07	.02	.01	.06	-.27
Working Memory						
Normalized Score	.13	.08	***.41***	-.01	-.21	.01
	Whole sample (*n* = 140)		Whole sample (*n* = 55)		Whole sample (*n* = 55)	
Max. Number of Digits recalled	**.23***		.07		-.18	

Note. AGL = Artificial Grammar Learning Task, ***italic =*** $|r| \geq .30$, gray = $BF_{10} \geq 3$.

* corrected $p < .05$.

[a] missing data of $n = 1$ six-year-old for these subtests.

[b] The result pattern for German Grammar II with AGL performance remained the same when excluding $n = 3$ adults with very low scores in this subscale ($T < 15$).

As a control analysis, 5-year-olds were excluded for correlations within Year 1 (i.e., associations between Learning Gains & language/memory skills, see Table 23 left column), because of their lack of learning effects in these first three sessions. This did not change the overall pattern of results. However, in this analysis, working memory capacities were positively associated with performance improvement across the first three sessions with the original stimulus set (AGL Session 3 – Session 1 ~ Working Memory (normalized score): $r = .35$, $p = .035$, $BF_{10} = 8.84$). This association, now evident for 6- and 7-year-olds, might have been occluded by the lack of performance improvements in 5-year-olds across the first three sessions.

Thus, within Year 1 a role of working memory in multi-session sequence learning emerged for the whole sample of all child groups and adults combined: Being able to immediately recall a longer digit span was associated to larger AGL performance improvements over three sessions. In the same vein, higher working memory capacities within the group of 7-year-olds showed a trend for being linked to stronger transfer of rule knowledge from stimulus set 1 to the new surface features of stimulus set 2.

3.2.2. Associations between cognitive skills in Year 1 and AGL performance in Year 2

To test if previously assessed memory and language skills (Year 1) are associated with later visual sequence (re)learning (Year 2), we calculated correlations of cognitive variables from the beginning of the study (Session 1 in Year 1) with AGL task performance across the one-year delay. For AGL task performance, we used four performance scores that quantified consolidation from Year 1 to Year 2 (Session 4 – Session 3), performance gains across three more sessions with stimulus material 1 in Year 2 (Session 6 – Session 4), and two transfer effects to stimulus material 2 in Year 2 (Transfer Savings: Transfer 2 – Session 4; Transfer Loss: Transfer 2 – Session 6; for details see *Data Analysis*). Again, these performance parameters were chosen to keep in line with the previous chapters that compared age groups in their long-term AGL (Chapters II & III).

Table 24 shows all correlations for these four AGL performance parameters in Year 2 with German grammar skills, declarative working memory and working memory scores in Year 1 (see Appendix D for all Scatterplots & *BF*s). A higher working memory capacity, as assessed for all age groups combined (longest digit span recalled), was associated with less AGL improvement over another set of three sessions with the first stimulus set (Session 6 – Session 4, $r = -.22$, $BF_{10} = 3.05$)[10] and a less pronounced drop in performance from the last session with the first stimulus set to the transfer session with stimulus set 2 (Transfer 2 – Session 6, $r = .23$, $BF_{10} = 2.76$). Both associations displayed *P*-values at the a-priori defined .05 level for statistical significance when corrected for multiple comparisons (both corrected $p = .052$). These associations are similar to what has been reported within Year 1, but they display the influence of working memory capacity (independent of age norms)

[10] To follow up on this negative direction of the link, an additional correlation between absolute performance levels in Session 4 and working memory was calculated ($r = .70$, $p < .001$, $BF_{10} > 100$): Persons with higher working memory skills in Year 1 started out at a better performance level in Year 2 (i.e., Session 4) than those with lower working memory. This additional result will be used in the discussion to evaluate the negative correlation of working memory with AGL (Session 6 – Session 4).

across a longer timescale, i.e., how memory capacity contributed to AGL performance after the one-year delay. These associations with working memory (normalized scores) did not emerge in separate correlations within the groups of children and adults (see Table 24).

From separate correlations in children and adults, which considered their cognitive scores normalized for age, moderate associations (n.s., but > |.30|) emerged only in the adult group[11]: Their declarative memory retrieval skills in Year 1 showed similar associations to AGL performance in Year 2, comparable to the role of working memory capacity in the whole sample; better memory retrieval in adults in Year 1 was associated with less AGL improvement with the first stimulus set (Session 6 – Session 4; $r = -.40$, $BF_{10} = 4.88$) and preserved performance in Year 2 in the second stimulus set as compared to the first stimulus set, i.e., less Transfer Loss (Transfer 2 – Session 6; $r = .35$, $BF_{10} = 2.22$). On a similar note, the less pronounced drop in generalization performance in Year 2, was associated with better (general) German grammar skills in Year 1 for the adult group (Transfer 2 – Session 6; $r = .33$, $BF_{10} = 1.60$). Despite these consistent trends across different cognitive skills displaying moderate associations on a descriptive level, these correlations with AGL performance in adults failed to reach statistical significance after correcting for multiple comparisons (both corrected $p \geq .050$).

In children, none of the associations between cognitive skills in Year 1 and AGL performance in Year 2 exceeded a value of |.30| or reached statistical significance (all $|r_s| \leq .21$, all corrected $p \geq .579$, all $BF_{10} \leq .88$)[12].

[11] One adult group completed an additional transfer session in the end of Year 1 (Adults 1, see Table 1B). Separately calculating all associations across one year for the two adult groups (cognitive skills in Year 1 and AGL in Year 2), yielded slightly different result patterns to those reported for both groups combined. These separate associations for adult groups are detailed in Appendix D.

[12] Excluding 7-year-olds for analyses between cognitive skills in Year 1 and AGL in Year 2, because of their additional transfer session in the end of Year 1, yielded the same pattern of results.

Table 24

Correlations for Memory & Grammar Skills in Year 1 with AGL Performance in Year 2

	Retention Year 1 to Year 2 AGL (Session 4 – Session 3)		**Learning Gains AGL** (Session 6 – Session 4)		**Transfer Savings AGL** (Transfer 2 – Session 4)		**Transfer Loss AGL** (Transfer 2 – Session 6)	
	Children (n = 63)	Adults (n = 40)	Children (n = 63)	Adults (n = 40)	Children (n = 63)	Adults (n = 40)	Children (n = 63)	Adults (n = 40)
German Grammar I (Plural)	-.17	.07	.09	-.04	.17	.10	.08	.15
German Grammar II (General)	-.21	.10 [a]	.01	-.05 [a]	.15	.26 [a]	.11	***.33***
Declarative Memory								
Encoding	-.07	.02	.09	-.08	-.01	.06	-.12	.15
Retrieval	.08	.15	.04	**-.40**[ƚ]	-.04	-.07	-.15	***.35***
Working Memory								
Normalized Score	-.02	.29	.07	-.10	.19	.10	.20	.21
	Whole sample (n = 103)		Whole sample (n = 103)		Whole sample (n = 103)		Whole sample (n = 103)	
Max. Number of Digits recalled	-.07		**-.22**[ƚ]		-.09		**.23**[ƚ]	

Note. AGL = Artificial Grammar Learning Task, ***italic*** **=** $|r| \geq .30$, gray = $BF_{10} \geq 3$.

[ƚ] corrected p = .05.

[a] The result pattern for German Grammar II with AGL performance remained the same when excluding n = 1 adult with very low scores in this subscale (T < 15).

Summarizing correlations from this section on how memory and language skills from Year 1 were associated with later visual sequence (re)learning in Year 2, we observed that a higher working memory capacity across all age groups was associated with smaller relearning gains across three more sessions with stimulus set 1, but with preserved performance from stimulus set 1 to stimulus set 2. Similar trends emerged within the adult group, especially for better memory retrieval skills in Year 1: Higher abilities to retrieve learned associations across a 30-minute delay in the beginning of the study were associated with the same AGL performance parameters in Year 2, i.e., with less additional improvement in stimulus set 1, but with stronger preserved transfer performance for stimulus set 2.

3.3. Are memory and grammar skills in Year 2 associated with multi-session sequence (re)learning performance after the delay?

We additionally tested a possible association of memory and language skills with visual sequence (re)learning within Year 2, i.e., when assessing all variables at the same timepoint after the one-year delay. For AGL task performance in Year 2, we used the same performance scores as in the previous section to quantify consolidation (Session 4 – Session 3), performance gains (Session 6 – Session 4) and transfer effects (Transfer Savings: Transfer 2 – Session 4; Transfer Loss: Transfer 2 – Session 6; for details see *Data Analysis*). As pointed out earlier (see *Assessment of memory and language skills*), analyses within Year 2 could only be conducted for the three age groups with available psychometric assessments for cognitive skills in Year 2 (5-year-olds, 6-year-olds and Adults 2).

Table 25

Correlations for Memory & Grammar Skills in Year 2 with AGL Performance in Year 2

	Consolidation Year 1 to Year 2 AGL (Session 4 – Session 3)		**Learning Gains AGL** (Session 6 – Session 4)		**Transfer Savings AGL** (Transfer 2 – Session 4)		**Transfer Loss AGL** (Transfer 2 – Session 6)	
	Children (*n* = 43)	Adults 2 (*n* = 18)	Children (*n* = 43)	Adults 2 (*n* = 18)	Children (*n* = 43)	Adults 2 (*n* = 18)	Children (*n* = 43)	Adults 2 (*n* = 18)
German Grammar I (Plural)	***.30***	-.11	.01	-.07	-.01	-.20	-.03	.16
German Grammar II (General)	-.05	.09	.12	.09	.01	***.37***	-.11	***.39***
Declarative Memory								
Encoding	.04	.09	.03	-.23	-.06	.24	-.09	**.68***
Retrieval	-.03	.06	.07	***.30***	.00	**.66***	-.07	***.50***
Working Memory								
Normalized Score	.03 [a]	-.14	.08	.04	.09	-.01	.01	.17
	Whole sample (*n* = 61)		Whole sample (*n* = 61)		Whole sample (*n* = 61)		Whole sample (*n* = 61)	
Max. Number of Digits recalled	-.02		-.14		-.07		.19	

Note. AGL = Artificial Grammar Learning Task, ***italic*** = | r | ≥ .30, gray = $BF_{10} \geq 3$.
* corrected $p < .05$.

Table 25 shows all correlations for the four AGL performance parameters with German grammar skills, declarative working memory and working memory scores in Year 2 (see Appendix D for all Scatterplots & *BF*s).

Within the adult group, statistically significant associations emerged for declarative memory skills and AGL transfer effects: Better memory retrieval in Year 2 was associated with more Transfer Savings across one week, i.e. a larger performance gain in the transfer session using stimulus set 2 (Transfer 2) compared to the first session of relearning with stimulus set 1 (Session 4, $r_{s\,=}$ .66, p = .015, BF_{10} = 18.10). At the same time, better memory encoding in Year 2 was associated with a smaller drop in transfer performance when compared to the last relearning session with stimulus set 1 (Transfer Loss: Transfer 2 – Session 6, $r_{s\,=}$ .68, p = .010, BF_{10} = 25.75).

This pattern reoccurred in associations for the same AGL performance parameter in association to memory retrieval and German grammar skills in adults ($r \geq |.30|$, but corrected $p > .05$): Stronger memory retrieval ($r_{s\,=}$ .50, BF_{10} = 3.08) and better German grammar skills (general, $r_{s\,=}$ .39, BF_{10} = 1.16) in Year 2 were correlated with a less pronounced Transfer Loss in Year 2 (Transfer 2 – Session 6). Better performance in German grammar tests in Year 2 furthermore showed a moderate association with gains in transfer performance (Transfer 2), as compared to the first session (Session 4) of relearning with stimulus set 1 (Transfer Savings; $r_{s\,=}$ .37, BF_{10} = 1.36). Apart from these memory and language associations with AGL transfer performance in adults, their memory retrieval in Year 2 also seemed to be related to an additional performance improvement for stimulus set 1 in relearning of Year 2 ($r_{s\,=}$ .30, BF_{10} = 1.24). All of the above memory and grammar skill associations to AGL performance within Year 2 showed consistent trends and moderate effect sizes on a descriptive level, but failed to reach statistical significance when corrected for multiple comparisons (all corrected $p \geq .088$).

In the children's group, a trend emerged for better (plural) grammar skills in Year 2 being positively related to consolidation of AG knowledge across the one-year delay (from Session 3 to Session 4; $r_{=}$ .30, BF_{10} = 1.18). This association did not reach statistical significance after correcting for multiple comparisons (corrected p = .260), however.

AGL associations with working memory in the whole sample, as measured by the max. length of correctly recalled digits independent of age, were not found to be statistically significant within Year 2 (all $|r_s| \leq .19$, all corrected p $\geq .602$, all $BF_{10} \leq .77$).

To sum up, associations in Year 2 emerged predominantly within the adult group, and therein mainly for declarative memory being related to two measures of AGL transfer performance. Some additional trends showed for German grammar skills, which were positively associated with AGL transfer performance in adults and consolidation across one year in children.

4. Discussion

The aim of Project 3 was to test if multi-session sequence learning in a visual AGL task is associated with general memory and language grammar skills in 5-year-olds, 6-year-olds, 7-year-olds and adults (1) across one week within Year 1, (2) across a one-year delay, and (3) across one week after the delay (Year 2) in an exploratory manner.

Before evaluating these associations, we confirmed that test-retest estimates across one year and parallel test estimates between the two presented stimulus sets were reasonably high for the used AGL task in our combined sample of children and adults. Both types of estimates ranged from r = .77 to r = .84, which is in line with the size of internal reliability coefficients reported in previous studies with adult and developmental samples (Farkas et al., 2023; Qi et al., 2019; Siegelman et al., 2017; Torkildsen et al., 2019). Thus, our AGL measures seem suitable for a correlational approach with memory and language outcomes in the sample investigated here.

AGL performance was associated by around r = .20 with working memory capacity across one week (see (1)) and one year (see (2)) in our whole sample, and with memory encoding and retrieval skills by size r = .50-.70 after the delay (see (3)) in adults. I will in turn discuss these exploratory findings and additional trends in our data, including possible contributions of German grammar skills to AGL performance in Year 2.

4.1. The role of working memory in multi-session sequence learning

In line with our hypotheses, memory skills were consistently associated with sequence learning outcomes on all three investigated timescales, i.e. across one week (Year 1), across one year (memory in Year 1 to AGL in Year 2), and after the delay (Year 2).

More specifically, stronger working memory capacities, assessed in Year 1, were associated with greater performance improvement in the first stimulus set within Year 1, but less additional improvement in the same stimulus set after the 12-month delay, i.e., at relearning in Year 2, across all age groups combined. Despite the somewhat counterintuitive negative association across one year, i.e., less additional improvement after the delay when scoring high in a working memory task at the beginning of the study, a closer look at our data confirmed that working memory facilitated learning: persons with higher working memory skills in Year 1 started out at a better performance level in Year 2 (i.e., Session 4) than those with lower working memory, probably leaving less room to further improve in the three relearning sessions (from Session 4 to Session 6). The above findings on working memory substantiate the positive link between working memory and sequence learning reported in single-session (Arnon, 2019) and multi-session (Smalle, Page, et al., 2017) studies in healthy adults, which was somewhat stronger (r = .30-.60) than the correlations reported here (approx. r = .20). We extended previous reports by showing that this link is also observed in a sample including children[13], using the same working memory measure across all ages (max. number of digits immediately recalled). Presentation speed has been shown to influence visual sequence learning in children aged 5 to 12 years (Arciuli & Simpson, 2011), confirming working memory involvement in this age range. One reason for Smalle, Page, et al. (2017) reporting working memory associations in adults, but not in children, could be their choice of stimulus material: They matched sequence length to the working memory capacities of adults (sequences with 9 items), while we presented sequences of a max. length of seven items to enable successful and stable AGL learning effects in (young) children.

Additionally, higher working memory skills in Year 1 in our whole sample were associated with stronger transfer in AGL in Year 2 (r = .23). This positive, albeit statistically non-significant (corrected p = .052), link was replicated within the group of 7-year-olds for associations within Year 1 (r = .41, corrected p = .175), who had available transfer data with stimulus set 2 already in Year 1. These preliminary findings in a multi-session AGL setting corroborate findings by Hendricks et al. (2013) in adults, who demonstrated that working memory skills are important for generalizing encountered regularities to new input. Based on our results, it can be speculated that this link holds across longer timescales of one week and even one year, as well as for developmental samples. Hendricks et al. (2013) argued that transfer conditions might specifically involve working memory capacities as part of an explicit decision process that is based on generating and testing hypotheses of sequence rules (hypotheses generation model by Dulany et al., 1984).

---

[13] Further corroborating a role of working memory skills in children's sequence learning, working memory capacities (Year 1) were positively associated with performance improvement across the first three sessions (Year 1), when excluding 5-year-olds from the sample of children (r = .35, see *Results*). This association in the combined sample of 6- and 7-year-olds might have been occluded by the lack of any AGL performance improvement in 5-year-olds across these sessions.

Both findings of the current study on working memory contributions in a multi-session AGL setting speak to the debated role of awareness and attention for sequence learning, i.e., whether sequential regularities are extracted and used in an automatic/implicit vs. effortful/explicit manner (Conway, 2020; Daltrozzo & Conway, 2014; Nemeth et al., 2013). Conway (2020), based on an extensive review of experimental studies, put forward that the extent to which working memory capacities are involved in sequence learning depends on task complexity (higher for complex regularities and transfer conditions), the development of underlying learning/memory systems with age (adults > children), the measurement of learning outcomes (direct/test phase performance > indirect/online measures during exposure), and possibly the timepoint of assessment during prolonged learning (later > earlier in the time course of a task). With regard to neurocognitive mechanisms, an implicit system underlying the extraction and use of sequential regularities has been suggested to get support from optional explicit processes, including working memory and related attentional resources, which can be deployed to varying degrees (Batterink et al., 2015).

Thus, our task setting with several sessions of AGL, which exposed participants to a complex set of underlying sequence rules and provided continuous performance feedback in multiple test phases per session, possibly elicited the recruitment of additional explicit learning mechanisms in all age groups (see Chapters II & III for similar levels of explicit knowledge in children and adults, acquired over the course of this study). This might have contributed to the role of working memory in sequence learning of a mixed sample with children and adults in the current study.

4.2. Declarative memory skills and the generalization of sequential regularities, dependent on age

In accord with predictions from the literature on a developmental shift towards increasingly relying on the declarative memory system for learning in adulthood (Ambrus et al., 2020; Gualtieri & Finn, 2022; Poldrack & Packard, 2003; Smalle et al., 2022), AGL associations with declarative memory emerged in adults, but not in children (see Table 23-25, evidence from Bayes statistics for children is discussed below). The clearest picture emerged for better declarative memory skills benefitting rule generalization to a second set of visual stimuli, which we observed within Year 2 for both measures of transfer (r = .68: memory encoding ~ AGL Transfer Savings; r = .66: memory retrieval ~ AGL Transfer Loss). For one transfer marker, this association was replicated as a trend across the one year delay (r = .35: memory retrieval Year 1 ~ AGL Transfer Loss Year 2). Additionally, our data showed trends for declarative memory being related to AGL performance improvements in the original, first, stimulus set on both timescales (see Table 24 & 25 for a negative trend across one year (r = -.40), but a positive trend within Year 2 (r = .30)).

These exploratory findings on declarative memory skills influencing adults' performance in a multi-session AGL task corroborate studies on different maturational trajectories of a late matured declarative memory system, as opposed to an early matured procedural memory system (Finn et al., 2016; Meulemans et al., 1998; Parkin & Streete, 1988). The assumption that children seem to rely less on declarative memory for learning was possibly reflected in the lack of associations with declarative memory skills in the present study: The *BF* supports evidence for no association (H_0) in children for associations across one year (all *BFs* $\leq$ .32) and within Year 2 (all *BFs* $\leq$ .23), but remains inconclusive for associations within Year 1 (all *BFs* $\leq$ .77, *BFs* for single associations see Appendix D).

West et al. (2018) have argued that the terms explicit/declarative and implicit/procedural have been used interchangeably in the learning and memory literature. Adopting this view, our findings fit well with the age-dependent shift towards more explicit learning, "model-based" mechanisms, which have been put forward for sequence learning across development (Conway, 2020; Daltrozzo & Conway, 2014; Janacsek et al., 2012; Nemeth et al., 2013): More explicit learning mechanisms in adults, involving top-down controlled attention and memory resources, might be reflected in their reliance on declarative memory skills for multi-session sequence learning. Children, on the other hand, might have relied more strongly on their procedural memory resources, or "model-free", learning processes, as proposed by this shift. We did not include separate psychometric measures to index procedural/implicit memory in this study, making it impossible to directly dissociate procedural from declarative contributions to learning performance in children.

Nevertheless, others have shown that declarative memory functions (Finn et al., 2016; Gathercole, 1998; Juhász & Németh, 2018; Parkin & Streete, 1988) and underlying neurocognitive systems (Brod et al., 2013) mature more slowly than procedural memory functions. Adding to this, Friederici et al. (2013) manipulated the recruitment of brain areas which usually contribute to explicit (declarative) learning in adults by transcranial direct current stimulation of the dorsolateral PFC (BA 9) during the acquisition of statistical regularities. This caused adults to engage more implicit learning mechanisms, reflected in later event-related potentials indexing associative processes similar to those observed in infants, as opposed to more controlled processes (reflected in an early N400) shown as an adult "default" processing mode in the sham condition (Friederici et al., 2013).

Based on the above studies, observed associations of sequence learning performance with declarative memory skills in adults vs. no such associations in children in the current study can be speculated to indicate age-dependent changes in the recruitment of declarative/explicit vs. procedural/implicit memory systems, which might hold for long-term learning settings as well. Neuroimaging evidence is needed to further test the neural recruitment of both memory systems at different ages, keeping in mind that children's brains seem to be functionally less specialized compared to adults, e.g., with regard to semantic (corresponding to declarative/explicit mechanisms) and syntactic (corresponding to procedural/implicit) processing (Brauer & Friederici, 2007).

Our results furthermore extend the existing literature by providing preliminary evidence that declarative memory skills in adults might be involved particularly when transferring encountered regularities to new surface features, similar to what has been demonstrated for the role of working memory in rule generalization (Hendricks et al., 2013). This seems to be in line with research which related generalization, e.g., in category learning, to processes of reactivation and retrieval across several instances of learning (Vlach, 2014). Providing an explanation for how the brain might implement transfer-related memory processes, Y. Liu et al. (2019) measured sequential replay in the human hippocampus with MEG during rest. They observed that a previously learned sequence rule was applied to new items, as reflected in observed neural activity patterns that followed the learned position rule instead of the actually experienced input sequence. It can be speculated that in persons with better memory skills, these processes of replay, reactivation and retrieval work more efficiently. These speculations about neural underpinnings of the memory associations observed here need to be tested by future research.

In sum, our memory results on adult multi-session AGL are in line with the "extraction and integration framework" (Thiessen, 2017), and with accounts on "chunking" (Perruchet, 2019; Pothos, 2007), which describe sequence learning as a set of memory processes. Adopting such a view on sequence learning has consequences apart from the theoretical appreciation, which exact computations might underlie sequence learning (Christiansen, 2018; Perruchet & Pacton, 2006). It is furthermore relevant for predicting how sleep and forgetting rates influence age differences in retention and relearning of sequential regularities, and under which conditions learned regularities might be transferred to new situations and input (see Chapter I; reviewed in Forest et al., 2023). These aspects will be discussed comprehensively in Chapter V.

However, all interpretations on the role of declarative memory skills in adult multi-session AGL, especially those within Year 2, remain rather speculative and need to be addressed by future research, given our small sample of 18 adults with available language and memory assessments in Year 2. Since relationships between declarative memory skills and AGL performance consistently emerged across several timescales and AGL measures, we think that they nevertheless hint to a genuine relationship that is worth further investigations.

4.3. Possible contributions of grammar skills to multi-session sequence learning

With regard to German grammar skills, we report trends in our multi-session study, which suggest that grammar skills might benefit sequence relearning in adults and children (consistently across all three timescales $| r | > .30$, but not statistically significant).

For children, stronger consolidation across one year was associated with better grammar skills in Year 2. This fits well with reports of stronger retention of auditory sequences across up to 12 months for children, who had initially performed better in a different language measure, namely vocabulary knowledge (Smalle, Page, et al., 2017). The association across one year reported there is of similar magnitude ($r = .37$) as the one in our study ($r = .30$, $BF = 1.18$). The current study included children as young as age 5 years at the beginning of the study (vs. 8-9 year-olds in Smalle, Page, et al., 2017), performing a visual AGL task with non-linguistic stimulus material (vs. an auditory immediate recall task with syllables in Smalle, Page, et al., 2017). For unselected, i.e., non-clinical, developmental samples, there has been a controversy as to whether sequence learning and language are reliably related (Conway et al., 2019; Krishnan & Watkins, 2019; West et al., 2018, 2019). Smalle, Page, et al. (2017) and the present results stressed that in a multi-session setting, such long-term links between *retention* measures of sequence learning and natural language skills can be observed in an unselected sample of children age 5 to 8 years.

For adults, initially higher grammar skills in Year 1 were numerically associated with better generalization, measured as preserved AGL performance from the first to the second stimulus set in Year 2. Relatedly, within Year 2, adults who performed better in the German grammar assessments, showed stronger rule generalization in both measures of AGL transfer performance as well ($r = .33$ - $.39$, $BF = 1.16$ – 1.60). If these findings are replicated, they might extend the established link between language skills and sequence learning (Conway et al., 2007; Misyak et al., 2010; Misyak & Christiansen, 2012; Smith et al., 2015) to longer timescales, and to measures of rule generalization in unselected populations.

Note that for 7-year-olds within Year 1, this link between generalization and grammar skills showed an opposite trend, with better grammar skills being related to a more pronounced drop in performance for the new stimulus set compared to the last session of the first stimulus set. These preliminary findings are in line with what Ferman and Karni (2010) showed for acquiring phonological aspects of a sequence learning task with syllables: All age groups mastered the pronunciation of the new "words", but 12-year-olds and adults outperformed 8-year-old children in their phonological competence for the artificial "language" stimuli. Additionally, adults improved to a greater degree than 8-year-old children across several sessions in their pronunciation. Thus, (prior) language knowledge seems to be involved in sequence learning (as proposed in Forest et al., 2023). Potential age differences in the link between language and sequence learning outcomes might speak to the role of prior knowledge from natural language experiences in transferring regularities from environmental patterns to new input (Hickey et al., 2019; Siegelman et al., 2018).

Despite failing to reach statistical significance, the above associations in children and adults thus point to a potential role of language skills in long-term visual retention and transfer of sequential regularities. I think that this long-term perspective on language associations provides an important proof-of-concept tool for future research. It allows checking if the tasks and study designs researchers came up with actually capture what they were designed for: Do the sequence learning tasks successfully model language mechanisms and processes which unfold across a more extended time period than a single experimental session? This check is necessary to relate findings using these tasks to more general developmental concepts like sensitive phases in development, where e.g., the heightened sensitivity towards environmental patterns was argued to facilitate early language learning (Gualtieri & Finn, 2022; Janacsek et al., 2012; Werker & Hensch, 2015). Possible implications from all three dissertation projects for the concept of sensitive phases will be discussed in Chapter V.

4.4. Limitations and future directions

While small associations with grammar skills emerged consistently on different timescales of AGL in our study setting, they were weaker in size and more restricted to less measures of sequence performance for these language outcomes than for memory skills. This seems to contradict the tradition of sequence learning tasks, modeling processes like word or grammar learning (Erickson & Thiessen, 2015; Romberg & Saffran, 2010), and the considerable body of literature on natural language processes being involved in sequence learning behavior (Conway et al., 2007; Misyak et al., 2010; Misyak & Christiansen, 2012; Smith et al., 2015) and neural processing (Conway & Pisoni, 2008; Goranskaya et al., 2016; Skosnik et al., 2002). Looking at our choice of cognitive assessments, however, we covered a broader range of memory functions (working memory, declarative memory encoding & retrieval) compared to grammar as a single language domain (different aspects assessed in 2 tasks). This was done to deal with limited attentional resources of young children, by restricting assessments to those skills that were most strongly hypothesized to contribute to learning in the present AGL task. Nevertheless, this choice of assessments might have contributed to our pattern of findings. Besides, uncovering associations between grammar skills and AGL in the current study, particularly for children**,** might have been difficult due to additional task and sample characteristics: First, we used non-linguistic, visual AGL task material which by nature might draw less on language skills than phonological tasks. Domain specific vs. general mechanisms in sequence learning (Arnon, 2019; Conway, 2020; Frost et al., 2015; Pavlidou & Bogaerts, 2019; Siegelman et al., 2018) and their developmental trajectories (Forest et al., 2023; Raviv & Arnon, 2017; Shufaniya & Arnon, 2018) are still debated. Consequently, observing (high) associations with verbal language skills might be less likely for visual/non-linguistic tasks than for auditory/linguistic tasks, especially in developmental samples. Secondly, a large number of children scored very high in our grammar assessments (see Scatterplots in Appendix D), limiting variance in these tasks in our high-performing sample. Limited variance in one variable, in turn, restricts how strongly variables can be correlated (reviewed, e.g., in Carretta & Ree, 2022).

It is important to keep in mind, however, that memory skills, in particular when assessed with verbal measures as in our study, are at least moderately associated (Kaufman et al., 2010) with language measures (see West et al., 2018 for a comprehensive model on memory and language measures in a large sample of 7-8-year-old children & Kaufman et al., 2010 for academic language achievement in a large sample of 16-17-year-old students). So, we refrain from pitting respective contributions of memory vs. language skills against each other in the current study. The preliminary findings for both domains documented here at least speak against a clear distinction for age-dependent contributions of (working) memory vs. language skills to sequence learning. This distinction might have been derived from the respective correlational patterns reported in Smalle, Page, et al. (2017): Working memory capacity was associated with sequence learning outcomes in adults but not in children, while language (vocabulary) skills correlated with sequence learning in children but not in adults. However, language and memory systems in the brain have been shown to substantially overlap (Ullman, 2004), which lead to the declarative/procedural model (Ullman, 2001; 2004): It proposes shared neurocognitive mechanisms of procedural memory with language grammar learning, and of declarative memory with vocabulary/word learning processes, respectively. Based on this neurocognitive model and the behavioral evidence in developmental samples from above (Kaufman et al., 2010; West et al., 2018), we suspect that a well-powered study on individual differences in multi-session sequence learning would paint a more nuanced picture for both, memory and language contributions, at different ages. These contributions are possibly influenced by the chosen sequence learning task, either modeling grammar (e.g., AGL tasks) or word-form (e.g., Hebb learning tasks) learning.

In general, extrapolating the exploratory results of this chapter warrants caution, as our study had limited power to look into individual differences (especially in our small adult sample with cognitive assessments in Year 2, see above). Given some methodological differences between the age groups included in this study, we checked if our reported result patterns would hold when (1) excluding 5-year-olds for analyses within Year 1, because of their lack of learning effects in these first three sessions, and when (2) excluding 7-year-olds for analyses between cognitive skills in Year 1 and AGL in Year 2, because of their additional transfer session in the end of Year 1 (analyses within Year 2 by default excluded this age group due to their study design, see *Methods*). Both control analyses yielded the same overall result pattern as the previous analyses. Correlations in adults across one year, however, yielded a slightly different result pattern, when calculated for the two groups separately compared to when collapsing across these groups, as reported above (cognitive skills Year 1 ~ AGL Year 2; Adults 1: additional transfer session in Year 1 [Transfer 1] vs. Adults 2: only one transfer session in the end of Year 2 [Transfer 2], see Table 1). These separate correlations for the two adult groups can be found in Appendix D (Table D.2).

4.5. Conclusion

To conclude, we report exploratory evidence for memory skills, and to a lower degree potentially for language grammar skills, being implicated in multi-session sequence learning outcomes. This was shown using a visual AGL task in 5-year-olds, 6-year-olds, 7-year-olds and adults, who displayed associations with these cognitive skills across one week, a one-year delay, and at relearning after the delay. In particular, memory skills were consistently associated with sequence learning improvements and rule transfer to a second set of visual surface features. These findings highlight how (1) working memory functions across all included ages and (2) declarative memory skills in adults influence the long-term extraction and use of sequential regularities from the environment. The exploratory findings of this study need to be confirmed by future research, but are well in line with the "extraction and integration framework" (Thiessen, 2017). This framework views sequence learning as being closely related to memory processes like chunking, (re)activating, integrating and retrieving pieces of information.

Chapter V: General Discussion

The present dissertation investigated how the developmental timing of several instances of sequence learning influences learning outcomes in the long run. Three child groups (5-year-olds, 6-year-olds, 7-year-olds) and adults learned visual sequences involving complex rules in a modified AGL task, across several sessions (see Table 1). This longitudinal design spanned a one-year delay between two sets of sessions (see Fig. 1 & Fig. 8) and tested multi-session learning across one week (Year 1), and multi-session relearning of the previously acquired rules with the same stimuli after 12 months (Year 2). Additionally, it assessed rule generalization to new visual surface features in a separate session and exploratory associations of sequence learning outcomes with memory and language skills.

The main findings of this dissertation entailed that children from 6 years onwards successfully learn complex visual sequence rules across several sessions (see Project 1 & Project 2). They use their acquired rule knowledge after a 12-month delay for quicker and additive relearning of the same input compared to Year 1, and for transfer to new but related input, both in an adult-like fashion. We observed successful transfer in the very last session in all age groups. Seven-year-olds and an adult group with the same study design, who both completed a transfer session in the end of each year, showed successful transfer already in Year 1 (see Project 1). While this study did not confirm that an earlier developmental timing of several learning instances results in better outcomes in the long run, it corroborates that prior learning results in quicker re-acquisition of sequence rules after a delay when controlling for unspecific maturational effects in children (see Project 2). Regarding more general cognitive skills involved in multi-session sequence learning, working memory and declarative memory encoding/retrieval were consistently associated with AGL task performance (see Project 3). These associations emerged across one week, a one-year delay and at relearning after the delay. Numerically smaller associations emerged for language grammar skills and sequence learning performance, which did not reach statistical significance.

In the following, I will first discuss findings from Year 1 on how children of age 5 to 7 years and adults compare in their learning outcomes within one session and across several subsequent sessions (over one-week's time), with regard to possible neural adaptations and neurocognitive mechanisms underlying the observed behavior. The respective cognitive make-up and neural infrastructure available at different points in life are suggested to interact with the requirements of a learning situation (e.g., acquiring vs. retaining/relearning sequential regularities). Relatedly, age-dependent mechanisms are likely constrained by how rule knowledge is represented in memory (reviewed in Forest et al., 2023), as formed from encountering sequential patterns in the environment. Theoretical implications from findings on relearning in another set of sessions in Year 2 after the long-term delay will be discussed separately below (see *Relearning after a long-term delay: Savings in learning, plasticity & sensitive phases*). Generalization effects in the current study and how they add to the literature of learning transfer will be discussed in a separate section as well (see *Generalization of rule knowledge: mechanisms and timescales*).

1. Mechanisms of multi-session learning in development

Successful sequence learning of the AG rules was observed in adults and in children from age 6 years onwards, i.e., for 6-year-olds and 7-year-olds in both years and for 5-year-olds in Year 2. Adults outperformed all child groups at any time and displayed a steeper learning curve. This effect was replicated in two independent adult groups (Adults 1, Adults 2). In Year 1 adults showed learning effects at an earlier timepoint, i.e., with less task exposure needed, in addition to performing on an overall higher level across all three sessions. This behavioral advantage with older age was also observed within childhood, in that a state-space model (Smith et al., 2005) identified a numerically earlier timepoint for stable above-chance performance in the very first session of Year 1 in 7-year-olds (30th test trial) vs. 6-year-olds (41st test trial). From this model, no stable above-chance performance was identified for the youngest group of 5-year-olds by the end of Session 3, i.e., after a total of 150 test trials and exposure to 270 grammatical sequences. Between the groups of 6-year-olds, 7-year-olds and adults, however, learning gains across three subsequent sessions in Year 1were indistinguishable. This suggests adult-like learning efficiency in children as young as 6 years old for building on acquired rule knowledge in subsequent learning instances across 1 week. Although, children might need more exposure to rule-abiding input at a younger age to successfully acquire these sequence rules in the first place. Additionally, adults and 7-year-old children in the present study did not differ significantly in their reported levels of explicit sequence knowledge after four sessions of learning, i.e. by the end of Year 1. This can be seen as indirect evidence that explicit processes contributed to a similar degree to multi-session AGL performance in both age groups[14].

[14] Note that the groups of 5-year-olds, 6-year-olds, and Adults 2 only completed a single assessment of explicit sequence knowledge in the end of Year 2, which will be discussed below (see *Generalization of rule knowledge: mechanisms and timescales*).

1.1. Learning-induced plasticity in the cortex and top-down control

Age-dependent constraints on sequence learning can be expected to play out at least on two levels: (1) Bottom-up tuning of perceptual networks and (2) top-down control by prefrontal areas (discussed in Conway, 2020). Conway (2020) merged several previous theories (Kral & Eggermont, 2007; P. J. Reber, 2013; White et al., 2013) to derive two main mechanisms, which he suggested for learning-induced adaptations in the cortex, in response to sequential input. (1) Perceptual circuits are gradually tuned to process sequential input more efficiently. This allows for the extraction of regularities during exposure and results in less neural effort required for rule-following input in the recruited network (i.e., visual networks for the visual AGL task investigated here). This mechanism has been elaborated by the “plasticity of processing” approach, which was established for implicit learning (P. J. Reber, 2013). Learning-induced plasticity, in this view, is incremental and distributed across the cortex depending on the task at hand but limited to the recruited network for respective task processing. Across development, Conway (2020) proposes that downstream information from higher order areas like the PFC increasingly influences the tuning of perceptual areas (2). The PFC is able to operate on longer timescales than perceptual areas, i.e., it can cover larger temporal receptive windows of input. Consequently, its influence allows for top-down control of the input features that are attended to and enables the integration of sequential input over time (see also Forest et al., 2023 for developmental considerations on the role of the inferior frontal gyrus in sequence learning).

Within this plasticity framework, the gradual tuning of task-specific visual circuits from 6 years onwards might underpin successful learning of complex visual regularities (“plasticity of processing”, P. J. Reber, 2013), as observed in the current study. General principles of cortical plasticity in sensory networks underlying learning were shown to be functional early in life already (McClelland et al., 1995). At the same time, earlier within-session rule acquisition and overall higher performance accuracy, as observed in older children and adults of the present study, can be speculated to reflect age-dependent changes in prefrontal control. As top-down regulation and neural substrates mature profoundly in pre-school and school-age children (reviewed in Bunge & Crone, 2009; Ramscar & Gitcho, 2007), this might have enabled more effective cortical adaptations in response to sequential input in older children and adults of the current study. Given that the group of 5-year-olds in Year 1 failed to perform above chance in our AGL task, it could be hypothesized that, to a certain extent, prefrontal control guiding learning-induced tuning of perceptual circuits was necessary to enable successful learning behavior in the current task setting. However, at age 5 years, young children might have been yet unable to exert the required prefrontal control. Indeed, Bunge and Zelazo (2016) have proposed that the ability to acquire and use increasingly more complex rules across childhood mirrors the maturational trajectory of specific PFC subregions: An early matured orbitofrontal cortex might enable young children to acquire simple associations between stimulus-response patterns, while successfully using higher-order rules might require a more mature (rostro)lateral PFC, which was shown to develop more slowly. Given the complex nature of the AG rule set in the current study, protracted PFC development, especially in lateral subregions (BA 9,10, 44-47; Bunge & Zelazo, 2016), and their role in rule learning might have contributed to the observed age pattern in learning behavior. In addition to these maturational effects, increased brain activation in regions of cognitive control has been reported to result from one year of schooling by tracking 5-year-olds who entered vs. such who did not enter first grade during that time (Brod et al., 2017). Thus, changes in prefrontal control due to maturation (children simply growing older) likely interacted with neural changes induced by environmental demands (e.g., children entering school) to first give rise to *successful* rule acquisition (around age 6), and then to *better* overall rule application with increasing age in our multi-session sequence task.

1.2. Age-differences in sequence learning: The implicit vs. explicit model

One influential model of sequence learning in development has put forward that the balance of implicit vs. explicit mechanisms recruited in learning situations shifts towards a more explicit learning mode (Conway, 2020; Daltrozzo & Conway, 2014; Janacsek & Nemeth, 2012; Nemeth et al., 2013), possibly around the age of 12 years (Janacsek et al., 2012; Nemeth et al., 2013). This shift was argued to reflect the maturational timeline of brain regions that underly more cognitively controlled, goal-directed learning, in particular with regard to the interplay of the hippocampus and the PFC (Janacsek et al., 2012). As they mature, these regions were suggested to increasingly provide internal models as interpretations of encountered events (“model-based” or supervised learning mode), taking over from a basic pattern “extraction” system that involves the striatum and basal ganglia (“model-free” or unsupervised learning mode; Janacsek et al., 2012; Nemeth et al., 2013). Driven by the latter brain regions which mature early, the implicit learning mode is thought to prevail at a younger age. Relying more on this implicit system, as (younger) children seem to do naturally, in turn was proposed to facilitate learning outcomes in certain situations. This early advantage was especially observed in contexts where less attentional control was beneficial for learning (e.g., picking up task-irrelevant information see Tandoc et al., 2022), and for indirect learning markers, like an implicit tracking of regularities "online" (Janacsek et al., 2012; Nemeth et al., 2013; Smalle, Page, et al., 2017).

The current study is not suited to evaluate how these two systems were recruited on the neural level. Yet, the main model proposition of a *more implicit* extraction of sequential regularities early in development that results in *better learning outcomes* than in adulthood, could not be confirmed in the present sample of children age 5 to 7 years and adults.[15] Thomas et al. (2004) showed that recruited brain circuits, comprising fronto-striatal networks, substantially overlap in adults and 7-11-year-old children during implicit tracking (finger-tapping) of a visuomotor sequence. Despite some activation differences in the hippocampus and motor areas between age groups, recruitment of these "implicit" networks during an implicit learning situation enabled better learning performance in adults as compared to children. Friederici et al. (2013) directly manipulated the recruitment of implicit vs. explicit mechanisms during sequential language learning in adults. They applied inhibitory stimulation over the dorsolateral PFC (BA 9), which resulted in implicit neurocognitive mechanisms as indexed by infant-like ERPs at test. However, adults who showed implicit mechanisms performed equally well at test, after learning sequential regularities from language input as compared to when using their "default" explicit processing mode in the sham condition (see, however, Smalle, Panouilleres, et al., 2017 ; Smalle et al., 2022). This challenges the notion that implicit learning mechanisms, which are taken to prevail "naturally" in younger learners, are only available early in life to support successful, or even superior, extraction of regularities and application of sequence knowledge. These observations are corroborated by H. Liu et al. (2023), who showed that implicit and explicit memory traces emerge simultaneously in adult sequence learning (see also Batterink et al., 2015; Conway, 2020), and can be dissociated by using indirect (implicit) vs. direct (explicit) learning markers. This implies that implicit vs. explicit processes, resulting in respective representations of sequence knowledge, are not employed in an either-or-fashion that is solely determined by development. Rather, an implicit learning advantage early in

[15] However, rule awareness was reported at a later timepoint during the study by young vs. older learners, see Chapter III. AGL task associations with declarative memory skills emerged only in adults, see Chapter IV. Both findings are in line with a developmental shift towards more explicit learning mechanisms, which has been discussed in the respective chapters.

development can be taken to mean broader input to young learners (put forward by Forest et al., 2023), which does not necessarily lead to better learning outcomes. Age-dependent constraints of the learner furthermore interact with situational demands, with an option for deploying explicit mechanisms as an additional resource (Batterink et al., 2015, reviewed in Conway, 2020).

Adopting the view that we mainly tapped explicit processes and representations with our direct learning marker, a post-test after exposure (as proposed by Forest et al., 2023; H. Liu et al., 2023), our findings of successful learning from 6 years onwards underscore the role of external factors for eliciting explicit learning mechanisms. These factors likely included rather complex regularities (multiple sequence rules defined by an AG, including non-adjacent dependencies), a high amount of exposure to regularities, and performance feedback at test when applying sequence knowledge (based on study features influencing the role of attention & working memory as reviewed in Conway, 2020; discussed before in Chapter IV). In such learning situations, explicit neurocognitive mechanisms might to some extent be already available in early to middle childhood (see, e.g., performance associations with working memory in Chapter IV). As demonstrated in our study, this could be the case for children as young as age 6 years, which is a lot earlier than the age around 12 years that was proposed for a "natural" developmental shift to more explicit mechanisms by Nemeth et al. (2013). Using this idea of promoting explicit learning mechanisms in children, studies have started to look into intervention techniques along this implicit-explicit dimension that might benefit e.g., specific language outcomes in clinical populations (see Baron & Arbel, 2022 for a framework in developmental language disorders). While previous literature has mainly stressed the advantages of an implicit learning mode for extracting regularities in situations when a broad attention focus is beneficial (e.g., for learning task-irrelevant information; Rohlf et al., 2017; Tandoc et al., 2022), explicit learning mechanisms might favor other aspects of learning like rule generalization (see, e.g., Ferman & Karni, 2014; H. Liu et al., 2023).

1.3. Memory representations: How does the acquisition vs. short-term retention of sequence knowledge depend on age?

The above discussions underscore that a broader perspective is needed when evaluating how sequence learning changes across development. This perspective should acknowledge that age constrains learning on many different levels that go beyond the implicit-explicit framework. For instance, development has been proposed to shape what kind of learning output is generated from being exposed to sequential regularities, i.e. what is represented in memory (Forest et al., 2023). Age-dependent rates of forgetting, sleep-dependent consolidation mechanisms, and processes of memory encoding, integration, and retrieval interact to form these representations (discussed in Chapter I based on Forest et al., 2023). In that respect, the main contribution of the current study to developmental sequence learning models lies in extending the investigated timescale to several sessions, assessed in early to middle childhood (3 groups of children age 5, 6, and 7 years). The present finding of age-invariant learning rates across subsequent sessions over one week aligns well with previous work on retained sequence knowledge across short delays (Tóth-Fáber et al., 2023; reviewed in Janacsek & Nemeth, 2012; Lerner & Gluck, 2019; discussed in Chapters II & III). This work has established that children (from age 6 years onwards) retain sequence knowledge across several hours to several days (Tóth-Fáber et al., 2023; Juhász & Németh, 2018; Savion-Lemieux et al., 2009 Ferman & Karni, 2010; Smalle, Page, et al., 2017). Evaluating how consolidation rates change depend on age yielded mixed results, but studies with a similar study design as ours (several sessions, a task which included non-deterministic rules & task feedback; Ferman & Karni, 2010, 2014), or a comprehensive approach with large samples and results supported by Bayes statistics (Tóth-Fáber et al., 2023), respectively, report that this knowledge is consolidated equally well in children age 7 years and older and adults across such short time periods.

Our study shows that memory traces of acquired complex rule knowledge are used successfully in subsequent encounters with the same input from 6 years of age onwards. In that respect, it seems vital to dissociate age-dependent capabilities to *acquire* sequence rules from capabilities to successfully *tap* existing *representations,* formed during acquisition (i.e., *consolidation* & *retention*). At acquisition, encountered regularities are extracted and encoded into memory, forming initially fragile representations (Walker, 2005). Consolidation then stabilizes these representations in an offline period without practice, storing them into long-term memory for later retrieval (Walker, 2005). At retrieval, successful retention (i.e., the same performance level as before the delay) or even improved performance (i.e., a higher performance level as before the delay) is taken to reflect consolidated rule knowledge (Tóth-Fáber et al., 2023; Walker, 2005). By reporting age-invariant short-term consolidation across 24 hours, Tóth-Fáber et al. (2023) showed that the developmental timeline established for the *acquisition* of sequence knowledge in a visuomotor task (best before age 12 years Janacsek et al., 2012) does not translate into an equivalent timeline for age-dependent *retention* (age-invariant from 7-76 years of age), using the very same task. This implies that age constraints play out differently for learning on different timescales.

A similar dissociation was found in clinical populations with a diagnosis of Dyslexia (Bogaerts et al., 2015) and Tourette syndrome (Tóth-Fáber, Tárnok, et al., 2021), who were reported to differ from controls in their acquisition capacities of sequence rules, but not in their retention thereof. Intriguingly, adults diagnosed with Dyslexia have been shown to need more practice, i.e., repeated learning of sequential input, to acquire sequence rules, and still perform worse at immediate recall even when provided with additional practice (Bogaerts et al., 2015; using an auditory Hebb learning task). Despite their slower and impaired acquisition, however, they retained their acquired sequence knowledge to the same degree as control participants across a one-month delay. This finding seems to mirror how age groups of the current study compare in their overall learning capacity vs. in their use of sequential regularities across several sessions: While children needed more exposure to rule-following input to successfully acquire complex sequence rules and overall performed worse in applying these rules than adults, they used their acquired rule knowledge in an adult-like fashion for subsequent learning. This was observed across short delays of, on average, 2-4 days between sessions, over one-week's time (long-term retention and relearning will be discussed later in section *3. Relearning after a long-term delay: Savings in learning, plasticity & sensitive phases*).

In an attempt to explain why sequence knowledge might be consolidated in an age-invariant manner despite age-dependent acquisition capacities, opposed neural dynamics depending on the learning phase have been suggested. The recruited brain areas were proposed to interact differently when *acquiring* new sequence rules vs. when drawing on the acquired rule knowledge for *retention*, as indexed by a competitive (acquisition) as compared to a cooperative (retention) dynamic (Tóth-Fáber et al., 2023). Adopting a sequence learning model from the motor domain (Albouy et al., 2013), Tóth-Fáber et al. (2023) suggested that a network involving the striatum, the hippocampus and the PFC collaboratively supports behavioral retention effects. At the core of this network, the striatum has been shown to be involved in memory retrieval (Scimeca & Badre, 2012; see also Chapter IV for correlations between memory retrieval and AGL performance in adults) and to mature early in development, underpinning sequence learning (Forest et al., 2023). In contrast, during the acquisition of sequential regularities, "implicit" striatal learning (supporting the extraction of basic probabilities) has been postulated to compete with "explicit" mechanisms based on hippocampus-PFC-interactions (generating internal models of the environment; Conway, 2020; Janacsek et al., 2012). Tóth-Fáber et al. (2023) argued that their proposed framework should hold for learning of temporal sequence rules that comprise non-adjacent dependencies, both of which applies to the AGL task used in the current study. While in need of neuroimaging evidence for direct support, this proposed dissociation in interactions related to memory processes might explain the present result pattern from multi-session sequence learning. It might be speculated that a striatum-based, collaborative, retention network is available quite early to children (possibly around age 6), allowing for the successful use of sequence knowledge across several instances of learning. At the same time, acquisition abilities, indexing rather "explicit" mechanisms based on the hippocampus/PFC in our task setting, seem to continuously develop across the age range investigated here.

2. Generalization of rule knowledge: Mechanisms and timescales

We did not observe greater rule transfer to new visual surface features in (younger) children, which had been predicted from stronger (over)generalization early in development. Instead, all investigated age groups transferred the learned rule knowledge to a new stimulus set to a similar degree on two timescales: This was the case for transfer after the first set of three sessions in Year 1 compared between 7-year-olds and Adults 1 (across 1 week, see Table 1 & Fig. 1), as well as for transfer after the second set of sessions in Year 2 compared between 5-year-olds, 6-year-olds and Adults 2 (across 1 year, see Table 1 & Fig. 8). The two groups who completed a transfer session in the end of each year (7-year-olds & Adults 1 from Project 1) did not show any additional benefit for the second stimulus set in Year 2 (Transfer 2, see Table 1) from another set of relearning sessions with the first stimulus set (Session 4-6, see Table 1). This generalized use of rule knowledge after the long-term delay was replicated in the other age groups (5-year-olds, 6-year-olds and Adults 2 from Project 2), who did not perform significantly better in the transfer session in Year 2 (Transfer 2, see Table 1) compared to the first session of relearning in Year 2 (Session 4).

Yet, age might have affected at which point during the study participants became aware of sequence rules in the task, reflected in children's and adults' responses to open questions: A larger number of 6-year-olds and adults compared to 5-year-olds reported to have noticed sequence rules in the AGL already at the end of Session 3 in Year 1 (similar to the age pattern in Smalle, Page, et al., 2017), which was the first assessment in the study (Project 2). Additionally, levels of knowledge about specific sequence rules (legal items in salient positions, legal item-item transitions), did not differ significantly between adults and children in this first assessment of Year 1 (Project 1). However, 7-year-olds explicit knowledge about specific rules (legal items in salient positions, legal item-item transitions) improved from Year 1 to Year 2[16]. Additionally, a positive link between explicit rule knowledge and rule transfer in the AGL task emerged in 7-year-olds in Year 1. After the delay, in the end of Year 2, levels of reported rule knowledge did not differ between 5-year-olds, 6-year-olds and Adults 2, or 7-year-olds and Adults 1, respectively.

These findings on transfer effects and verbalizable rule knowledge have been discussed in previous chapters (see Chapter II & III), with regard to possible underlying brain mechanisms and study characteristics which might have promoted transfer, like spreading out learning instances across time. Summarizing the main aspects from these discussions, sequential replay, involving the hippocampus and cortex, could allow for regularities to be extracted and the resulting cortical representations to be strengthened (Janacsek & Nemeth, 2012; Lerner & Gluck, 2019; Y. Liu et al., 2019) – a mechanism that has been closely related to sleep-dependent consolidation (Lerner & Gluck, 2019). The consolidation of sequence knowledge in an offline period with sleep, in turn, has been proposed to be driven by becoming aware of underlying sequence rules (Janacsek & Nemeth, 2012). Additional factors, like extensive rule exposure before testing transfer, performance feedback at test, and offline periods between sessions promoting “abstraction-by-forgetting” (Vlach, 2014), have been proposed to foster generalization, particularly in young populations as observed in the present study for children aged 5 years at first exposure (discussed in detail in Chapter III).

Converging findings of all age groups within the current project, the following discussion will focus on how development might influence the transfer of sequential regularities by zooming in on mechanisms of generalization. A recent review article by Taylor et al. (2021) dissociates two generalization accounts. First, generalization can be viewed to result from integrated representations, formed from binding at encoding. Second, generalization can be based on “on-the-fly” processes of simultaneously reactivating separate representations at retrieval. Based on this distinction, two lines of evidence on (1) age-dependent properties of *memory representations* and on (2) age-related changes in *memory*

[16] This was not the case for adults in the same study design (Adults 1). Seven-year-olds and Adults 1 completed two assessments of explicit knowledge about specific rules, since they participated in an additional home follow-up in Year 2 that was originally not planned for. Specific knowledge of sequence rules in all other age groups was only assessed in the end of Year 2, to avoid inducing rule awareness or any change in strategy for the second set of sessions after the delay.

processes (e.g., encoding & retrieval) need to be reconciled (see also Chapter I). Additionally, the different timescales of the current project will be considered, which included testing generalization across several subsequent sessions over one week, and after two sets of sessions spanning a one-year delay (see sessions for all age groups in Table 1).

Our age-invariant transfer effects in children age 5-7 years and adults did not reflect what has been proposed for age-related changes in memory representations formed from sequence learning: Forest et al. (2021) suggested that representations become increasingly specific in childhood up to roughly age 7, before turning into broader representations that include specific *and* additional higher-order information (discussed as fuzzy, specific and broad representations in Forest et al., 2023). Yet, a major difference in terms of the two approaches by Taylor et al. (2021) which was summarized above is that Forest et al. (2021) investigated generalization on the level of *integrated representations* (transitional probabilities between adjacent items making up triplets, stored in “specific” representations & chunks of triplet items grouped independently of their order, stored in “broad” representations). Whereas in our task, we did not measure transfer for integrated input that had been presented close in time before (see “broad” representations above). Rather, underlying sequence rules had to be applied to completely new input at transfer (rule-based generalization in Taylor et al., 2021), making it impossible to bind any perceptual features from the original input into a joint representation that could be used for generalization (similarity-based generalization in Taylor et al., 2021). This means that in our task, transfer was likely to always involve “on-the-fly” generalization processes at retrieval (see Taylor et al., 2021). In this view, generalization maps onto the “extraction and integration framework” (Thiessen, 2017), which models sequence learning as a set of memory processes like chunking, (re)activating, integrating and retrieving pieces of information (discussed in Chapter IV). In line with this, AGL task performance, including measures of transfer, was associated with working memory capacity in the whole sample and with declarative memory encoding/retrieval in adults (see Chapter IV).

Taylor et al. (2021) caution against inferring the specific properties of representations from observed generalization behavior, but put forward that generalization mechanisms might change over time. First, successful generalization could be based on reactivating separate representations simultaneously at retrieval, which then are step-by-step converted into a generalized, integrated representation by repeated reactivation. Additionally, cognitive control and prior knowledge were proposed to influence how information from several learning experiences are integrated into memory across development (Brod et al., 2013), possibly from late (making inferences at retrieval) in middle childhood to early (forming integrated representations at encoding) memory stages in adulthood (Shing et al., 2019). According to Keresztes et al. (2018), the balance of generalization mechanisms which draw on memory encoding and retrieval shifts across childhood: Driven by protracted hippocampus development, decreasing processes of pattern completion alongside increasing processes of pattern separation lead to more specific encoding and retrieval around the age of 6 years (elaborated in Chapter I). These age-related changes in memory might have put (younger) children at an advantage for generalization based on "on-the-fly" reactivation of separate representations at retrieval (Taylor et al., 2021) in the current study. Our observation of children age 5 to 7 years and adults transferring learned regularities to the same degree to new input, can thus be argued to result from interactions between two age-dependent generalization mechanisms, involving (1) integrated representations and (2) on-the-fly retrieval processes (see Taylor et al., 2021). As elaborated above, representations and memory processes implicated in generalization have been suggested to show opposing developmental patterns in the age range investigated here.

Our task design likely favored abstraction processes at all ages due to two reasons (see Taylor et al., 2021 for additional task-related influences on generalization): (1) Periods of sleep were interspersed between sessions before testing transfer (Lerner & Gluck, 2019; Y. Liu et al., 2019), and (2) direct learning markers were applied that presumably tapped explicit memory traces (Forest et al., 2023), which were shown to grow more abstract over time (H. Liu et al., 2023). To observe successful, age-invariant transfer from age 7 onwards, three subsequent sessions of rule exposure with the same stimulus set across one week seem to have sufficed. [17] This can be concluded from 6-year-olds in Year 2 performing on the same level as age-matched controls at transfer, despite 6-year-olds' previous rule exposure in 3 additional sessions of Year 1 (Project 2), and from no additional transfer gains compared to Year 1 in 7-year-olds and Adults 1 after 3 sessions of relearning in Year 2 (Project 3, see also discussions on ceiling effects in adults in Chapter II; see Chapter III for considerations on what might be required for younger children to show transfer effects).

---

[17] For additional evidence in support of this interpretation, see comparisons between the two adult groups in their transfer performance (Adults 1: across 1 week & after 3 sessions with stimulus set 1, Adults 2: across 1 year & after 6 sessions with stimulus set 1) in Appendix C.

To sum up, age seems to constrain a learner's generalization abilities (termed "perceptual and cognitive biases" in Aslin & Newport, 2012), in addition to external factors, i.e., influences related to the input properties and learning context. Both types of influences likely contributed to the observed transfer effects in the multi-session study on sequence learning. Adult-like rule transfer to a completely new stimulus category in children as young as 5 years at first exposure, as demonstrated here, emphasizes children's early abilities to flexibly use encountered patterns from the environment. Additionally, the present findings assign an important role to situational factors that determine if children generalize what they have learned to new input and situations. The current results furthermore stress that qualitative over quantitative changes in sequence learning across childhood deserve to be investigated more closely (see Forest et al., 2023).

3. Relearning after a long-term delay: Savings in learning, plasticity & sensitive phases

Our results showed retained rule knowledge in 6-year-olds, 7-year-olds and adults, who started out at their final performance level reached one year earlier. Despite failing to perform above-chance in all three sessions of Year 1, the group of 5-year-olds improved across three more sessions after the 12-month delay (Year 2) to the same degree as older age groups (6-year-olds, 7-year-olds, adults). This means that all age groups showed similar relearning rates with the original stimulus set further boosting their final performance level in Year 2 compared to Year 1. Additionally, relearning benefits emerged as within-session learning effects for all groups of children: They needed less (re)exposure to sequence rules to perform above chance in Year 2 compared to Year 1. When controlling for unspecific (maturational) effects in Year 2 in both child groups with naïve age-matched controls available (5-year-olds and 6-year-olds), evidence emerged for a genuine effect of prior learning on their performance after a one-year delay. This mainly showed in earlier learning effects of 5-year-olds and 6-year-olds in the first session of Year 2 than when being first exposed to these sequence rules: Modeling performance on a trial basis, both groups exceeded chance performance at least two task blocks earlier than naïve children of the same age. Remarkably, the group of 5-year-olds displayed this faster relearning in Year 2 due to prior task exposure despite the fact that they had failed to show learning of sequence rules in any behavioral markers of Year 1. Nevertheless, these relearning analyses did not confirm that an earlier developmental timing of several AGL instances results in better learning outcomes in the long run.

These findings on relearning in children of age 6 (5-year-olds in Year 2), 7 (6-year-olds in Year 2), 8 (7-year-olds in Year 2) years and adults offer a new perspective on the long-term use of complex sequence rules across development, which to our knowledge has not been described for several relearning instances after a long-term delay before. Characterizing relearning after a long-term delay at different ages helps to refine the concept of lasting memory traces from previous experiences, adding to the understanding of plasticity processes and sensitive phases in development.

3.1. (How) do long-term benefits from prior sequence learning change across development?

By demonstrating facilitated relearning across extended time periods (up to one year), this dissertation provides proof-of-concept that adults and children aged 5 to 7 years use previously acquired rule knowledge for several additional learning instances in visual sequence learning after a long-term delay. Remaining sequence knowledge was tested in this study for the first time (1) for children younger than 8 years after a long-term delay, and (2) with regard to its role for relearning across another set of multiple sessions instead of a single follow-up session. In the group of 5-year-olds, earlier relearning effects observed due to prior learning (defined as “savings” in learning by Ebbinghaus, 1880; elaborated in Chapter I) demonstrate that encountering environmental patterns might leave stable, yet temporarily hidden or dormant memory traces, which can be reactivated later on to support behavioral advantages.

Behavioral savings have been shown already in infants as young as 5 to 6 months, after a delay of 48 hours (Cornell, 1979). They showed that pre-exposed infants needed less than the minimal presentation time usually required by naïve infants to show subsequent picture recognition. Parkin and Streete (1988) compared three groups of children age 3 to 7 years (3-year-olds, 5-year-olds and 7-year-olds) and adults, all in the same behavioral recognition task which was administered after one-hour as well as after a two-week delay. They reported comparable, baseline-controlled savings for all age groups in recognizing familiar vs. new fragmented pictures, i.e., identifying a picture at a more fragmented stage with less perceptual information available (see, however, Parkin & Streete, 1988 for an age gradient with greater savings in young age for learning familiar vs. new picture pairs).

But what could be the underlying neural mechanisms for relearning behavior in the long run? Underlying behavioral savings in sensory and motor skills across a delay of several months in rodents, structural plasticity on the level of dendritic spines has been reported (Xu et al., 2009; Yang et al., 2009). These skill-specific spines were shown to emerge during the acquisition of such new skills and to be later reactivated after a delay, when being exposed to the same learning environment. Thus, this mechanism was suggested to enable faster and more efficient relearning behavior in trained vs. naïve animals (Xu et al., 2009; Yang et al., 2009). Hofer and Bonhoeffer (2010) proposed that these findings from non-human animal studies might explain relearning observations in humans as well, as first reported by Ebbinghaus (1880) for faster and more efficient learning of familiar vs. new syllable lists. Temporarily dormant adaptations in the task-specific neural infrastructure can be speculated to account for the result pattern observed in the group of 5-year-olds here: The first set of learning sessions before the delay can be hypothesized to have caused lasting, possibly dendritic, adaptations in the cortical circuits recruited by the sequence learning task. Respective neural "preparations" from before delay might have facilitated relearning after the delay when the group of 5-year-olds was re-exposed to the same sequential regularities – despite the fact that early stages of a learning-induced plasticity in the underlying infrastructure had not been captured in their learning behavior before the delay. Spanning a shorter time interval of 48 hours, Cornell (1979) reported that 6-month-old infants displayed a similar reactivation of seemingly forgotten knowledge, which could not be detected in behavior without re-exposure. This corroborates the observed phenomenon on a shorter timescale, in this case possibly driven by additional neural mechanisms of plasticity that allow for short-term, learning-induced adaptations. These short-term adaptations on a neural level are likely rather physiological than structural in nature (see *Learning-induced plasticity in the cortex and top-down control*).

Age-invariant relearning advantages after a long-term delay reported in the current study furthermore extend previous research outside the classical savings literature. Successful consolidation of sequential information, reflected in retained or even improved performance levels, has been reported in different age groups and across various delays of up to several weeks (reviewed in Janacsek & Nemeth, 2012; Lerner & Gluck, 2019). Our finding adds to existing literature on retention across 2-12 months by characterizing how previous rule knowledge is used in multiple additional sessions after a long-term delay, instead of in a single follow-up as in previous studies (Ferman & Karni, 2010, 2014; Kóbor et al., 2017; Smalle et al., 2017; Tóth-Fáber et al., 2021, discussed in Chapters II & III). The present study corroborates these findings, broadening the investigated age range to children as young as 5 years, and provides evidence for relearning effects after a one-year delay on two timescales:

(1) quicker acquisition within the first session and (2) additive learning leading to higher performance levels across another set of several sessions. Even when accounting for very high performance levels in adults, possibly limiting a child-adult comparison, age-invariant relearning patterns remain informative for children aged 5 to 7 years. Comparable relearning rates in this age range seem surprising, given that children's general cognitive skill set matures profoundly during that time (Sameroff & Haith, 1996, discussed in Chapter III). Additionally, the oldest children in the current study (7-year-olds) experienced most rule exposure before the delay, due to their additional transfer session in Year 1. The youngest children (5-year-olds) nevertheless benefitted from relearning in Year 2 to the same degree, which shows that they were able to use previously acquired rule knowledge efficiently for later learning already at a pre-school age (see also previous discussions on age-invariant retention despite age-related acquisition in *Memory representations: How does the acquisition vs. short-term retention of sequence knowledge depend on age?*).

While the above studies are in line with our observations, they contradict the idea that the influence of prior learning might be *greatest* early in development. Direct support of an relearning advantage at a young age comes from two studies. Livosky and Sugar (1992) observed greater savings in 3-year-olds (> 5-year-olds > young adults) after two weeks, measured as the number of repetitions needed to reach a learning criterion for familiar relative to new picture pairs. In a different study with older children (age 8-9 years), long-term retention of implicit sequence knowledge (i.e., without any explicit task component included), was reported to be better than in adults up to a delay of 12 months (Smalle, Page, et al., 2017). Children's performance after the delay even exceed the retention level predicted from their previous learning trajectories by a fitted power-law function, while adults' performance declined (Smalle, Page, et al., 2017). Based on these findings, one could expect that drawing on prior knowledge is most effective when acquired at a young age. This claim, however, has not been systematically investigated to our knowledge – neither with children of different ages, nor outside retrospective studies with natural language experiences (e.g., J. S. Johnson & Newport, 1989; Scherag et al., 2004; elaborated below in *A developmental timeline for how prior experiences are used in the long run has to be established to further inform sensitive periods*). The current study, however, does not provide support for age-dependent relearning advantages, at least in the investigated age range (5- to 7-year-old children and adults) and measured across the limited time of one year. How this finding might inform sensitive phases for extracting and using sequential regularities from the environment, will be discussed in the next paragraph.

3.2. Considering long-term plasticity and sensitive periods underlying learning

How do the present results on relearning after a long-term delay speak to the concept of sensitive phases in development? Sensitive periods have been characterized as times of heightened responsiveness to certain types of input that profoundly shape later learning, which are underpinned by greater neural plasticity (Knudsen, 2004). Originally, these time periods and underlying mechanisms of neural plasticity had been described for circumscribed functions, like ocular dominance in vision (reviewed in Röder et al., 2021) or phoneme discrimination in speech perception (reviewed in Werker & Hensch, 2015). In a broader attempt to explain childhood advantages in certain cognitive outcomes, however, this concept has been translated to underlying, rather domain-general mechanisms (Gualtieri & Finn, 2022; Janacsek et al., 2012). For extracting sequential patterns from the environment, a proposed sensitive phase up to age 12 years was mainly based on cross-sectional evidence in the single-session *acquisition* of visuomotor regularities, as reflected in reaction-time measures (Janacsek et al., 2012). If early in life, prior exposure to sequential regularities mattered to a greater degree for later encounters with the same input as well, one would expect a clear pattern of quicker relearning resulting in higher performance levels after a long-term delay in younger vs. older age groups. Yet, this more pronounced long-term impact of prior learning early in development, which we had also hypothesized to find in the current cohort study, failed to show in any of the relearning markers measured here. No influence of age emerged, neither in identifying the timepoint of first learning from within-session modeling, nor in learning rates across a second set of sessions with the original stimulus set (Session 4 to 6), nor in rule transfer to a second stimulus set (discussed separately before, see *Generalization of rule knowledge: mechanisms and timescales*).

What does this mean in the context of sensitive periods? Perhaps the present cohort design was not able to capture a more pronounced long-term impact of prior learning early in development, due to a limited time period of delay (one year) tested in a limited age range (3 child groups of age 5 years, 6 years and 7 years). With regard to longer delays possibly needed for an age-at-acquisition-effect to unfold, evidence from natural language experiences in populations with a history of emigration or adoption has provided some insight (e.g., J. S. Johnson & Newport, 1989; Scherag et al., 2004; elaborated below in *A developmental timeline for how prior experiences are used in the long run has to be established to further* *inform sensitive periods*). In experimental settings that controlled for the exact onset, amount and characteristics of sequential input, however, sequence knowledge in adults was retained for delays from one to three years – despite having been acquired already at an adult age (Allen & Reber, 1980; M. C. Frank et al., 2012; Romano et al., 2010). We are not aware of any similar experimental studies that included different age groups to test how different ages of learning onsets influence learning outcomes across extended delays of several years. A challenge that remains for future studies addressing this gap in childhood is to control for maturational vs. genuine experience-dependent effects after a very long-term delay. The age range of 4 to 7 years was chosen for the current study, since it has been identified as a period of change in learning mechanisms, e.g., from the memory and generalization literature (Keresztes et al., 2017). It has been suggested that neural plasticity reflecting the successful (neural) *acquisition* of (automatic) auditory rules presented in spoken language might be restricted to even earlier time windows until four years of age (Mueller et al., 2018). However, when testing the consequences for the long-term (behavioral) *retention* of auditory rules in children vs. adults, Smalle, Page, et al. (2017) still found a childhood advantage in children as old as 8-9-years vs. adults for an implicitly acquired syllable sequence (see discussions in *Memory representations: how does the acquisition vs. short-term retention of sequence knowledge depend on age?*). This challenges the proposition that the children chosen for the current study (age 5 to 7 years) were simply too old to display a childhood advantage over adults at relearning.

A methodological limitation of the current study that has been discussed before (Chapters II & III), are very high performance levels in adults from Session 2 in Year 1 onwards. For uncovering child-adult differences at relearning, however, this resulted in “more room to improve” (Juhasz et al., 2019) for children vs. adults in Year 2 (and for younger compared to older children, respectively). Nevertheless, children did not show higher relearning rates than adults.[18] Thus, it is unlikely that ceiling effects in adult performance occluded age differences in relearning that might have supported a sensitive period account of sequence learning. In contrast, this methodological constraint should have increased the likelihood of finding a childhood advantage in relearning, at least as reflected in

[18] One might argue, however, that a loss in knowledge across one year from the last session in Year 1 to the first session in Year 2 might have been less detectable in adults, due to their performance at ceiling in the end of Year 1.

steeper relearning curves across the second set of sessions in Year 2. This notwithstanding, the observed comparable performance gain in all groups of children across the second set of sessions after the delay demonstrates an impressive relearning capability already at a young age. Given that the applied AGL task was more difficult for younger children than for older children and adults, as reflected in 5-year-olds performing at chance in Year 1 despite extensive rule exposure and task practice, their behavior at relearning can be taken as evidence for a very effective use of prior sequence knowledge.

3.2.1. A developmental timeline for how prior experiences are used in the long run has to be established to further inform sensitive periods

It is also possible that the current findings point to theoretical aspects of sensitive periods, which need to be reconsidered. First, trying to identify sensitive periods for very broad behavioral categories (e.g., speech perception or sequence learning per se) could prove to be difficult and hardly meaningful. For instance, in speech perception, Werker and Hensch (2015) put forward that sensitive phases for increasingly more complex abilities (e.g., first for phoneme discrimination, then for phonological categories) build upon each other across several years of life to enable successful (native) language acquisition in the long run ("cascading nature" described in Werker & Hensch, 2015, fig 3). It is less clear, how exactly this postulated timeline of several consecutive sensitive periods underlying the successful acquisition of a certain function, might influence later learning outcomes that tap temporarily unused knowledge after a delay.

As mentioned earlier, studies on populations with different natural language experiences in childhood have only been able to contrast persons who had experienced vs. lacked language input until a certain age (reviewed in Werker & Hensch, 2015). Long-term effects from these diverging language experiences in childhood usually have been tested in adulthood after many years, sometimes even decades. In doing so, studies have established that some aspects of language acquisition, in particular syntactical functions (J. S. Johnson & Newport, 1989; Scherag et al., 2004), seem to be linked to restricted times in (early) childhood. While this research has provided important insights into timelines of natural language acquisition, it was not able to vary the onset of language input systematically, let alone the precise nature and amount of this input. Furthermore, interfering input, e.g., from learning an additional language while being deprived from one's native language, makes it impossible to test how specific and persistent prior knowledge is preserved (discussed in Werker & Hensch, 2015).

Thus, more research is needed to describe which developmental steps allow for an efficient, further *use* of acquired knowledge over childhood. Relatedly, Tóth-Fáber et al. (2023) identified a gap in the literature for how the short-term consolidation of sequence knowledge changes across development. They show that a developmental timeline established for the *acquisition* of visuomotor regularities does not necessarily translate into an equivalent timeline for its *retention* across 24 hours (see above in *Memory representations: How does the acquisition vs. short-term retention of sequence knowledge depend on age?*). The current study might provide a starting point for a more careful and extensive investigation of how the long-term use of prior rule knowledge changes with age. Future research should consider the different levels of age-related changes in sequence learning (reviewed in Forest et al., 2023), especially with regard to how the acquired information is represented in memory at different ages and how learning outcomes are operationalized (see, e.g., indirect vs. direct learning markers in Forest et al., 2023).

3.2.2. Determinants of long-term learning plasticity and implications for sensitive periods

The present findings on age-invariant relearning raise the question what savings in learning as an index of long-term plasticity can say about sensitive periods. An answer might entail a more nuanced perspective on interrelations between the three concepts of (1) learning savings in behavior, (2) plasticity processes underlying savings and determining (3) the timing of sensitive periods, which in turn are used to explain (age) differences in long-term learning outcomes. In this respect, adults benefitting to the same degree from prior learning as children makes a case for rather effective adult plasticity even outside the proposed sensitive period. A sensitive period was suggested to end by age 12 years for most efficiently extracting (simple) sequential regularities from the environment in an implicit manner (Janacsek et al., 2012). Related literature on language learning has postulated that most efficient rule processing and acquisition is restricted to even earlier time windows (J. S. Johnson & Newport, 1989; Mueller et al., 2018). These time windows have been proposed to already close by the age of 7 years for native-like grammar acquisition of a second language (J. S. Johnson & Newport, 1989). A respective age pattern did not show up in the current study, neither on a short timescale, across one week of sequence learning of children age 5 to 7 and adults (discussed in *Mechanisms of multi-session learning in development*), nor in their relearning of sequence rules after a one-year delay. With regard to plasticity processes underlying relearning behavior, additional factors, apart from development, might thus influence long-term adaptations which enable the successful reactivation of previously acquired sequence knowledge.

For one, long-term plasticity in response to patterned input might depend more on external factors, like extensive practice and attentional focus, than previously thought. Earlier studies in adults which assessed retained sequence knowledge across extended time periods of up to three years, are indeed in line with this proposition (Allen & Reber, 1980; M. C. Frank et al., 2012; Romano et al., 2010). Despite using different task paradigms of visuomotor (Romano et al., 2010), artificial grammar (Allen & Reber, 1980), and artificial language learning (M. C. Frank et al., 2012) in rather small samples to look into sequence learning, they all report successful retention of the previously acquired sequential input across very long timescales of one (Romano et al., 2010), two (Allen & Reber, 1980) and three (M. C. Frank et al., 2012) years. This adult long-term retention seemed to be driven by a large amount of exposure/practice (M. C. Frank et al., 2012; Romano et al., 2010) and attentional effects (Allen & Reber, 1980). In the study of Allen and Reber (1980), attended features of the sequence input (letter strings governed by an AG) were manipulated in two acquisition conditions, which prompted different strategies during exposure. These different acquisition strategies were still reflected in participant's response patterns after a two-year delay without any additional exposure. The authors took this finding to point to distinct memory representations, which depended on a participant's attentional focus before the delay. In some cases, such external factors might even outweigh established developmental constraints on learning plasticity. For instance, M. C. Frank et al. (2012) showed that very extensive exposure to a continuous stream of syllables (10 days with 1 hour of listening each) lead to adults successfully identifying the most frequent words after the delay, which had been embedded in the input stream. This successful retention was tested three full years after mere exposure.

This finding in adults seems surprising, since durable and extensive brain adaptations in response to mere exposure have been put forward as a distinctive characteristic of sensitive periods during development (Keuroghlian & Knudsen, 2007; Rohlf et al., 2017). However, several authors proposed that it might be possible to "re-open" sensitive periods (termed critical periods in this context) (Cisneros-Franco et al., 2020; Dehorter & Del Pino, 2020), for instance by focused attention and extensive practice (Werker & Hensch, 2015). Werker and Hensch (2015) argued that repeated training involving performance feedback and systematic scaffolding of input (easy to difficult) can result in, e.g., successful discrimination of nonnative consonants, even after the developmentally defined sensitive phase for this function has already passed (Bradlow et al., 1997; Lively et al., 1994; McCandliss et al., 2002). Reopening sensitive periods by training has been speculated to be mediated by inducing neural processes that shift the brain's excitatory-inhibitory balance, or remove molecular brakes through, e.g., epigenetic changes (Werker & Hensch, 2015).

On a process level, what might enable a person to successfully tap the lasting traces formed from previous learning? Respective cognitive mechanisms should be functional later in life, too, when large-scale plasticity processes in the brain might be more limited than early in development. Apart from beneficial effects of practice and attention that have been discussed above, Romano et al. (2010) suggest that the "procedural reinstatement theory" for skill retention can provide a suitable framework (Fendrich et al., 1991; Healy et al., 2004; Marmie & Healy, 1995). According to this theory, recruiting the same set of cognitive processes (alongside perceptual and motor procedures), which are "reinstated" when re-encountering a previous learning situation, is at the core of retaining a complex skill across extended time periods. This theory allows for cognitive strategies and neural underpinnings to differ between participants during task acquisition, as long as the same set of processes can be successfully reinstated at the point of relearning. The exact neurocognitive mechanisms, which supported AGL task performance in the current study, are likely to differ in the age range investigated here (see previous discussions in *Mechanisms of multi-session learning in development*). In this context, the "procedural reinstatement theory" points out that reapplying the same task and procedures after one year might have been sufficient to reactivate task specific processes, which had previously supported task performance in an individual, independent of age. Using latent class modeling on several data sets from previous studies, even adult participants have been reported to rely on different response strategies for classifying the very same sequences in AGL tasks, after having been exposed to the very same stimulus sets (Visser et al., 2009). This has been taken as evidence that they relied on different types of information in the same input for their decisions, possibly using different mechanisms for extracting regularities (computing transitional probabilities between adjacent items vs. chunking sets of adjacent items). These findings in adult samples, which have been substantiated elsewhere (Pothos, 2007; Siegelman et al., 2019), make a case for individual differences in processing and using sequential information that go beyond age influences.

To sum up, long-term plasticity processes seem to depend on additional factors apart from developmental constraints. Rather, the successful use of previously acquired knowledge can be influenced by manipulating characteristics of learning environments, such as training and focused attention. In addition, it might make sense to first characterize how prior learning experiences are used in the long run at different ages, before proposing sensitive phases for broad behavioral functions like a sensitivity for sequential regularities.

4. Concluding remarks

The findings of the present dissertation corroborate the idea that lasting memory traces are formed from encountering sequential regularities in the environment, which can be successfully tapped later on (“savings” in learning, Ebbinghaus, 1880). This has been demonstrated in a multi-session study using an AGL task with complex visual sequence rules, which showed that children age 5 to 7 years used their acquired rule knowledge after a 12-month delay for quicker relearning of the same input compared to before the delay, and improved across relearning sessions in an adult-like fashion. While this study did not confirm that younger children display better learning outcomes at any timescale (neither at initial learning, nor at relearning), prior learning enabled quicker re-acquisition of sequence rules after a delay even after controlling for unspecific maturational effects in children, as well as in young children for whom learning seemed to be not successful before the delay.

The finding that learning seems to change *qualitatively* between 5 and 6 years of age, i.e., children only showed learning effects at age 6, not at age 5, can be related to similar observations in neighboring domains, like perceptual learning. Fiser and Lengyel (2022) have proposed that perceptual learning, statistical learning (termed “sequence learning” throughout this dissertation), and more abstract rule learning in vision all rely on a common computational mechanism that can be modeled in a Hierarchical Bayesian Framework. Adopting this view of a common learning process tested on different abstraction levels, our findings fit well with studies from perceptual learning, which report great changes between age 5 and 6 years: Around age 6, (1) immediate cross-modal recalibration in response to spatially disparate audio-visual input (Rohlf et al., 2020), and (2) the automatic use of external reference frames for localizing tactile input (Pagel et al., 2009; Röder et al., 2014) were shown to first emerge (reviewed in Bruns & Röder, 2023). This implies a general developmental shift in learning around age 5 to 6 years across different domains (see also Del Giudice, 2014; Sameroff & Haith, 1996), which might be driven by age-dependent changes in cognitive control and in the way encountered information is represented in the brain (see, e.g., Gualtieri & Finn, 2022; M. H. Johnson & Munakata, 2005; Ramscar & Gitcho, 2007).

The present results are furthermore in line with the "extraction and integration framework" (Thiessen, 2017), which views sequence learning as being closely related to memory processes like chunking, (re)activating, integrating and retrieving pieces of information. This account is supported by associations of working memory capacity and declarative memory encoding/retrieval skills with AGL task performance on different timescales, shown here. The proposed memory framework is further in line with the notion that successful rule transfer to a new visual category observed in all our age groups can be argued to have relied on simultaneously reactivating separate memory representations at retrieval (Taylor et al., 2021).

Adding to theories of learning mechanisms and sensitive periods, the current study proposes that the reliance on implicit vs. explicit learning modes and long-term plasticity processes could be less constrained by development than previously thought. For instance, extensive training and focused attention seem to promote the successful use of previously acquired knowledge independent of age in the age range investigated here. This stresses the role of learning environments, which can be targeted by interventions to promote long-term learning success. How exactly prior learning is represented on a neural level and what is necessary for its effects to persist, warrants further investigation. Future research can build on non-human animal-models (Hofer & Bonhoeffer, 2010), which have reported persistent structural changes in the cortex in skill acquisition and relearning after a long-term delay.

References

Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, *23*(11), 985–1004. https://doi.org/10.1002/hipo.22183

Allen, R., & Reber, A. S. (1980). Very long term memory for tacit knowledge. *Cognition*, *8*(2), 175–185.

Ambrus, G. G., Vékony, T., Janacsek, K., Trimborn, A. B., Kovács, G., & Nemeth, D. (2020). When less is more: Enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *Journal of Memory and Language*, *114*, 104144. https://doi.org/10.1016/j.jml.2020.104144

Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science*, *14*(3), 464–473. https://doi.org/10.1111/j.1467-7687.2009.00937.x

Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, *36*(2), 286–304. https://doi.org/10.1111/j.1551-6709.2011.01200.x

Arciuli, J., & Torkildsen, J. v. K. (2012). Advancing Our Understanding of the Link between Statistical Learning and Language Acquisition: The Need for Longitudinal Data. *Frontiers in Psychology*, *3*, 324. https://doi.org/10.3389/fpsyg.2012.00324

Arnon, I. (2019). Do current statistical learning tasks capture stable individual differences in children? An investigation of task reliability across modality. *Behavior Research Methods.* Advance online publication. https://doi.org/10.3758/s13428-019-01205-5

Aslin, R. N. (2017). Statistical learning: A powerful mechanism that operates by mere exposure. *Wiley Interdisciplinary Reviews. Cognitive Science*, *8*(1-2). https://doi.org/10.1002/wcs.1373

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science*, *21*(3), 170–176. https://doi.org/10.1177/0963721412436806

Baron, L. S., & Arbel, Y. (2022). An Implicit-Explicit Framework for Intervention Methods in Developmental Language Disorder. *American Journal of Speech-Language Pathology*, *31*(4), 1557–1573. https://doi.org/10.1044/2022_AJSLP-21-00172

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Bauer, P. J., Wenner, J. A., Dropik, P. L., & Wewerka, S. S. (2000). Parameters of remembering and forgetting in the transition from infancy to early childhood. *Monographs of the Society for Research in Child Development*, *65*(4), i-vi, 1-204. https://pubmed.ncbi.nlm.nih.gov/12467092/

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Bogaerts, L., Szmalec, A., Hachmann, W. M., Page, M. P. A., & Duyck, W. (2015). Linking memory and language: Evidence for a serial-order learning impairment in dyslexia. *Research in Developmental Disabilities*, *43-44*, 106–122. https://doi.org/10.1016/j.ridd.2015.06.012

Bogaerts, L., Szmalec, A., Maeyer, M. de, Page, M. P. A., & Duyck, W. (2016). The involvement of long-term serial-order memory in reading development: A longitudinal study. *Journal of Experimental Child Psychology*, *145*, 139–156. https://doi.org/10.1016/j.jecp.2015.12.008

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: Iv. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, *101*(4), 2299–2310. https://doi.org/10.1121/1.418276

Brauer, J., & Friederici, A. D. (2007). Functional neural networks of semantic and syntactic processes in the developing brain. *Journal of Cognitive Neuroscience*, *19*(10), 1609–1623. https://doi.org/10.1162/jocn.2007.19.10.1609

Brod, G., Bunge, S. A., & Shing, Y. L. (2017). Does One Year of Schooling Improve Children's Cognitive Control and Alter Associated Brain Activation? *Psychological Science*, *28*(7), 967–978. https://doi.org/10.1177/0956797617699838

Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience*, *7*, 139. https://doi.org/10.3389/fnbeh.2013.00139

Bruns, P., & Röder, B. (2023). Development and experience-dependence of multisensory spatial processing. *Trends in Cognitive Sciences*, in press.

Bunge, S. A., & Crone, E. A. (2009). Neural correlates of the development of cognitive control. *Neuroscience Research*, *55*, 22–37. https://doi.org/10.1017/CBO9780511757402.005

Bunge, S. A., & Zelazo, P. D. (2016). A Brain-Based Account of the Development of Rule Use in Childhood. *Current Directions in Psychological Science*, *15*(3), 118–121. https://doi.org/10.1111/j.0963-7214.2006.00419.x

Calabro, F. J., Murty, V. P., Jalbrzikowski, M., Tervo-Clemmens, B., & Luna, B. (2020). Development of Hippocampal-Prefrontal Cortex Interactions through Adolescence. *Cerebral Cortex*, *30*(3), 1548–1558. https://doi.org/10.1093/cercor/bhz186

Carretta, T. R., & Ree, M. J. (2022). Correction for range restriction: Lessons from 20 research scenarios. *Military Psychology*, *34*(5), 551–569. https://doi.org/10.1080/08995605.2021.2022067

Choi, D., Batterink, L. J., Black, A. K., Paller, K. A., & Werker, J. F. (2020). Preverbal Infants Discover Statistical Word Patterns at Similar Rates as Adults: Evidence From Neural Entrainment. *Psychological Science*, 956797620933237. https://doi.org/10.1177/0956797620933237

Christiansen, M. H. (2018). Implicit Statistical Learning: A Tale of Two Literatures. *Topics in Cognitive Science.* Advance online publication. https://doi.org/10.1111/tops.12332

Cisneros-Franco, J. M., Voss, P., Thomas, M. E., & Villers-Sidani, E. de (2020). Critical periods of brain development. *Handbook of Clinical Neurology*, *173*, 75–88. https://doi.org/10.1016/B978-0-444-64150-2.00009-5

Conway, C. M. (2020). How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning. *Neuroscience and Biobehavioral Reviews*, *112*, 279–299. https://doi.org/10.1016/j.neubiorev.2020.01.032

Conway, C. M., Arciuli, J., Lum, J. A. G., & Ullman, M. T. (2019). Seeing problems that may not exist: A reply to West et al.'s (2018) questioning of the procedural deficit hypothesis. *Developmental Science*, *22*(4), e12814. https://doi.org/10.1111/desc.12814

Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*(3), 356–371. https://doi.org/10.1016/j.cognition.2009.10.009

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

Conway, C. M., & Pisoni, D. B. (2008). Neurocognitive basis of implicit learning of sequential structure and its relation to language processing. *Annals of the New York Academy of Sciences*, *1145*(1), 113–131. https://doi.org/10.1196/annals.1416.009

Cornell, E. H. (1979). Infants' recognition memory, forgetting, and savings. *Journal of Experimental Child Psychology*, *28*(2), 359–374. https://doi.org/10.1016/0022-0965(79)90095-X

Daltrozzo, J., & Conway, C. M. (2014). Neurocognitive mechanisms of statistical-sequential learning: What do event-related potentials tell us? *Frontiers in Human Neuroscience*, *8*(324), 100. https://doi.org/10.3389/fnhum.2014.00437

Danner, D., Hagemann, D., & Funke, J. (2017). Measuring Individual Differences in Implicit Learning with Artificial Grammar Learning Tasks. *Zeitschrift Für Psychologie*, *225*(1), 5–19. https://doi.org/10.1027/2151-2604/a000280

Dehorter, N., & Del Pino, I. (2020). Shifting Developmental Trajectories During Critical Periods of Brain Formation. *Frontiers in Cellular Neuroscience*, *14*, 283. https://doi.org/10.3389/fncel.2020.00283

Del Giudice, M. (2014). Middle Childhood: An Evolutionary-Developmental Synthesis. *Child Development Perspectives*, *8*(4), 193–200. https://doi.org/10.1111/cdep.12084

DeMaster, D., Coughlin, C., & Ghetti, S. (2016). Retrieval flexibility and reinstatement in the developing hippocampus. *Hippocampus*, *26*(4), 492–501. https://doi.org/10.1002/hipo.22538

Deocampo, J. A., Smith, G. N. L., Kronenberger, W. G., Pisoni, D. B., & Conway, C. M. (2018). The Role of Statistical Learning in Understanding and Treating Spoken Language Outcomes in Deaf Children With Cochlear Implants. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 723–739. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0138

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*, 781. https://doi.org/10.3389/fpsyg.2014.00781

Döpfner, M., Plück, J., & Kinnen, C. (2014). *CBCL/6-18R, TRF/6-18R, YSR/11-18R. Deutsche Schulalter- Formen der Child Behavior Checklist von Thomas M. Achenbach*. Göttingen: Hogrefe.

Dulany, D., Carlson, R. A., & Dewey, G. (1984). A case of syntactical learning and judgment: How conscious and how abstract? *Journal of Experimental Psychology: General*. https://www.semanticscholar.org/paper/A-case-of-syntactical-learning-and-judgment%3A-How-Dulany-Carlson/66f5df620c2d01739074647fd60c32754ae5c28b

Dunabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Ebbinghaus, H. (1880). *Urmanuskript "Ueber das Gedächtniß"*. Passau: Passavia Universitätsverlag.

Endress, A. D., & Johnson, S. P. (2021). When forgetting fosters learning: A neural network model for statistical learning. *Cognition*, *213*, 104621. https://doi.org/10.1016/j.cognition.2021.104621

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research : JSLHR*, *52*(2), 321–335. https://doi.org/10.1044/1092-4388(2009/07-0189)

Farkas, B. C., Krajcsi, A., Janacsek, K., & Nemeth, D. (2023). The complexity of measuring reliability in learning tasks: An illustration using the Alternating Serial Reaction Time Task. *Behavior Research Methods.* Advance online publication. https://doi.org/10.3758/s13428-022-02038-5

Fendrich, D. W., Healy, A. F., & Bourne, L. E. (1991). Long-term repetition effects for motoric and perceptual procedures. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *17*(1), 137–151. https://doi.org/10.1037//0278-7393.17.1.137

Ferman, S., & Karni, A. (2010). No Childhood Advantage in the Acquisition of Skill in Using an Artificial Language Rule. *PloS One*, *5*(10), e13648. https://doi.org/10.1371/journal.pone.0013648

Ferman, S., & Karni, A. (2014). Explicit versus implicit instruction: Which is preferable for learning an artificial morphological rule in children? *Folia Phoniatrica Et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *66*(1-2), 77–87. https://doi.org/10.1159/000363135

Finn, A. S., Kalra, P. B., Goetz, C., Leonard, J. A., Sheridan, M. A., & Gabrieli, J. D. E. (2016). Developmental dissociation between the maturation of procedural memory and declarative memory. *Journal of Experimental Child Psychology*, *142*, 212–220. https://doi.org/10.1016/j.jecp.2015.09.027

Finn, A. S., Lee, T., Kraus, A., & Hudson Kam, C. L. (2014). When it hurts (and helps) to try: The role of effort in language learning. *PloS One*, *9*(7), e101806. https://doi.org/10.1371/journal.pone.0101806

Fiser, J., & Lengyel, G. (2022). Statistical Learning in Vision. *Annual Review of Vision Science*, *8*, 265–290. https://doi.org/10.1146/annurev-vision-100720-103343

Fló, A., Benjamin, L., Palu, M., & Dehaene-Lambertz, G. (2022). Sleeping neonates track transitional probabilities in speech but only retain the first syllable of words. *Scientific Reports*, *12*(1), 4391. https://doi.org/10.1038/s41598-022-08411-w

Forest, T., Abolghasem, Z., Finn, A., & Schlichting, M. L. (2021). *Memories of structured input become increasingly distorted across development.* PsyArXiv. https://doi.org/10.31234/osf.io/7y6z9

Forest, T., Schlichting, M. L., Duncan, K. D., & Finn, A. S. (2023). Changes in statistical learning across development. *Nature Reviews Psychology.* Advance online publication. https://doi.org/10.1038/s44159-023-00157-0

Frank, M. C., Slemmer, J. A., Marcus, G. F., & Johnson, S. P. (2009). Information from multiple modalities helps 5-month-olds learn abstract rules. *Developmental Science*, *12*(4), 504–509. https://doi.org/10.1111/j.1467-7687.2008.00794.x

Frank, M. C., Tenenbaum, J. B., Gibson, E., & Snyder, J. (2012). Learning and Long-Term Retention of Large-Scale Artificial Languages. *PloS One*, *8*(1), e52500. https://doi.org/10.1371/journal.pone.0052500

Frank, S. M., Bründl, S., Frank, U. I., Sasaki, Y., Greenlee, M. W., & Watanabe, T. (2021). Fundamental Differences in Visual Perceptual Learning between Children and Adults. *Current Biology : CB*, *31*(2), 427-432.e5. https://doi.org/10.1016/j.cub.2020.10.047

Franke, G. H. (2017). *BSCL - Brief-Symptom-Checklist. Deutsche Fassung* (1. Auflage). Göttingen: Hogrefe.

Friederici, A. D., Mueller, J. L., & Oberecker, R. (2011). Precursors to natural grammar learning: Preliminary evidence from 4-month-old infants. *PloS One*, *6*(3), e17920. https://doi.org/10.1371/journal.pone.0017920

Friederici, A. D., Mueller, J. L., Sehm, B., & Ragert, P. (2013). Language learning without control: The role of the PFC. *Journal of Cognitive Neuroscience*, *25*(5), 814–821. https://doi.org/10.1162/jocn_a_00350

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus 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

Gathercole, S. E. (1998). The Development of Memory. *Journal of Child Psychology and Psychiatry*, *39*(1), 3–27. https://doi.org/10.1111/1469-7610.00301

Gogtay, N., Nugent, T. F., Herman, D. H., Ordonez, A., Greenstein, D., Hayashi, K. M., Clasen, L., Toga, A. W., Giedd, J. N., Rapoport, J. L., & Thompson, P. M. (2006). Dynamic mapping of normal human hippocampal development. *Hippocampus*, *16*(8), 664–672. https://doi.org/10.1002/hipo.20193

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

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gómez, R. L. (2017). Do infants retain the statistics of a statistical learning experience? Insights from a developmental cognitive neuroscience perspective. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *372*(1711). https://doi.org/10.1098/rstb.2016.0054

Goranskaya, D., Kreitewolf, J., Mueller, J. L., Friederici, A. D., & Hartwigsen, G. (2016). Fronto-Parietal Contributions to Phonological Processes in Successful Artificial Grammar Learning. *Frontiers in Human Neuroscience*, *10*, 551. https://doi.org/10.3389/fnhum.2016.00551

Gualtieri, S., & Finn, A. S. (2022). The Sweet Spot: When Children's Developing Abilities, Brains, and Knowledge Make Them Better Learners Than Adults. *Perspectives on*

Psychological Science : A Journal of the Association for Psychological Science, 17456916211045971. https://doi.org/10.1177/17456916211045971

Healy, A. F., Wohldmann, E. L., & Bourne, L. E. (2004). The Procedural Reinstatement Principle: Studies on Training, Retention, and Transfer. In A. F. Healy (Ed.), *Decade of Behavior. Experimental cognitive psychology and its applications* (pp. 59–71). American Psychological Association. https://doi.org/10.1037/10895-005

Hendricks, M. A., Conway, C. M., & Kellogg, R. T. (2013). Using dual-task methodology to dissociate automatic from nonautomatic processes involved in artificial grammar learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(5), 1491–1500. https://doi.org/10.1037/a0032974

Hickey, A. J., Hayiou-Thomas, M. E., & Mirković, J. (2019). Grammatical Generalisation in Statistical Learning: Is it Implicit and Invariant across Development? *Proceedings of the Annual Meeting of the Cognitive Science Society*, *41*, 1908–1914.

Hofer, S. B., & Bonhoeffer, T. (2010). Dendritic spines: The stuff that memories are made of? *Current Biology : CB*, *20*(4), R157-9. https://doi.org/10.1016/j.cub.2009.12.040

Horst, J. S. (2013). Context and repetition in word learning. *Frontiers in Psychology*, *4*, 149. https://doi.org/10.3389/fpsyg.2013.00149

Hübener, M., & Bonhoeffer, T. (2014). Neuronal plasticity: Beyond the critical period. *Cell*, *159*(4), 727–737. https://doi.org/10.1016/j.cell.2014.10.035

Hudson Kam, C. L., & Chang, A. (2009). Investigating the cause of language regularization in adults: Memory constraints or learning effects? *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *35*(3), 815–821. https://doi.org/10.1037/a0015097

Hughes, C. (2013). Chapter 24 - Executive Function: Development, Individual Differences, and Clinical Insights. In J. L. Rubenstein & P. Rakic (Eds.), *Neural Circuit Development and Function in the Brain* (pp. 429–445). Academic Press. https://doi.org/10.1016/B978-0-12-397267-5.00062-5

Jabès, A., & Nelson, C. A. (2015). 20 years after “The ontogeny of human memory: A cognitive neuroscience perspective,” where are we? *International Journal of Behavioral Development*, *39*(4), 293–303. https://doi.org/10.1177/0165025415575766

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

Janacsek, K., & Nemeth, D. (2012). Predicting the future: From implicit learning to consolidation. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, *83*(2), 213–221. https://doi.org/10.1016/j.ijpsycho.2011.11.012

JASP Team (2021). JASP (Version 0.14.1) [Computer Software].

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Johnson, M. H., & Munakata, Y. (2005). Processes of change in brain and cognitive development. *Trends in Cognitive Sciences*, *9*(3), 152–158. https://doi.org/10.1016/j.tics.2005.01.009

Jost, E., Conway, C., Purdy, J., & Hendricks, M. (2011). Neurophysiological Correlates of Visual Statistical Learning in Adults and Children. *Proceedings of the Cognitive Science Society*, *Vol. 33*(No. 33).

Juhasz, D., Nemeth, D., & Janacsek, K. (2019). Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between- and within-group differences in average response times. *PloS One*, *14*(7), e0215116. https://doi.org/10.1371/journal.pone.0215116

Juhász, D., & Németh, D. (2018). Age-related and gender differences in conslidation of implicit sequence learning between 7 and 29 years of age. *Psychiatria Hungarica : A Magyar Pszichiatriai Tarsasag tudomanyos folyoirata*, *33*(2), 125–137.

Jung, Y., Walther, D. B., & Finn, A. S. (2020). Children automatically abstract categorical regularities during statistical learning. *Developmental Science*, e13072. https://doi.org/10.1111/desc.13072

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

Keresztes, A., Bender, A. R., Bodammer, N. C., Lindenberger, U., Shing, Y. L., & Werkle-Bergner, M. (2017). Hippocampal maturity promotes memory distinctiveness in childhood and adolescence. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(34), 9212–9217. https://doi.org/10.1073/pnas.1710654114

Keresztes, A., Ngo, C. T., Lindenberger, U., Werkle-Bergner, M., & Newcombe, N. S. (2018). Hippocampal Maturation Drives Memory from Generalization to Specificity. *Trends in Cognitive Sciences*, *22*(8), 676–686. https://doi.org/10.1016/j.tics.2018.05.004

Keuroghlian, A. S., & Knudsen, E. I. (2007). Adaptive auditory plasticity in developing and adult animals. *Progress in Neurobiology*, *82*(3), 109–121. https://doi.org/10.1016/j.pneurobio.2007.03.005

Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*(1), 171–184. https://doi.org/10.1037/a0025405

Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, *16*(8), 1412–1425. https://doi.org/10.1162/0898929042304796

Kóbor, A., Janacsek, K., Takács, Á., & Nemeth, D. (2017). Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Scientific Reports*, *7*(1), 760. https://doi.org/10.1038/s41598-017-00807-3

Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews*, *56*(1), 259–269. https://doi.org/10.1016/j.brainresrev.2007.07.021

Krishnan, S., & Watkins, K. E. (2019). A challenge for the procedural deficit hypothesis: How should we measure sequential learning in childhood? *Developmental Science*, *22*(4), e12815. https://doi.org/10.1111/desc.12815

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology*, *3*, 598. https://doi.org/10.3389/fpsyg.2012.00598

Kudo, N., Nonaka, Y., Mizuno, N., Mizuno, K., & Okanoya, K. (2011). On-line statistical segmentation of a non-speech auditory stream in neonates as demonstrated by event-related brain potentials. *Developmental Science*, *14*(5), 1100–1106. https://doi.org/10.1111/j.1467-7687.2011.01056.x

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What Learning Systems do Intelligent Agents Need? Complementary Learning Systems Theory Updated. *Trends in Cognitive Sciences*, *20*(7), 512–534. https://doi.org/10.1016/j.tics.2016.05.004

Langnes, E., Sneve, M. H., Sederevicius, D., Amlien, I. K., Walhovd, K. B., & Fjell, A. M. (2020). Anterior and posterior hippocampus macro- and microstructure across the lifespan in relation to memory-A longitudinal study. *Hippocampus*, *30*(7), 678–692. https://doi.org/10.1002/hipo.23189

Lavenex, P., & Banta Lavenex, P. (2013). Building hippocampal circuits to learn and remember: Insights into the development of human memory. *Behavioural Brain Research*, *254*, 8–21. https://doi.org/10.1016/j.bbr.2013.02.007

Lawrence. M. A. (2016). ez: Easy Analysis and Visualization of Factorial Experiments. R package version 4.4-0. https://CRAN.R-project.org/package=ez

Lee, J. K., Fandakova, Y., Johnson, E. G., Cohen, N. J., Bunge, S. A., & Ghetti, S. (2020). Changes in anterior and posterior hippocampus differentially predict item-space, item-time, and item-item memory improvement. *Developmental Cognitive Neuroscience*, *41*, 100741. https://doi.org/10.1016/j.dcn.2019.100741

Lenth, R. V, (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.6.1. https://CRAN.R-project.org/package=emmeans

Leon Guerrero, S., Smith, S. & Luk, G. (2016). Home language usage and cognitive control in bilingual preschoolers. In J. W. Schwieter (Ed.), *Cognitive control and consequences of multilingualism* (Bilingual processing and acquisition, volume 2). Amsterdam: John Benjamins Publishing Company.

Lerner, I., & Gluck, M. A. (2019). Sleep and the extraction of hidden regularities: A systematic review and the importance of temporal rules. *Sleep Medicine Reviews*, *47*, 39–50. https://doi.org/10.1016/j.smrv.2019.05.004

Liu, H., Forest, T. A., Duncan, K., & Finn, A. S. (2023). What sticks after statistical learning: The persistence of implicit versus explicit memory traces. *Cognition*, *236*, 105439. https://doi.org/10.1016/j.cognition.2023.105439

Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. J. (2019). Human Replay Spontaneously Reorganizes Experience. *Cell.* Advance online publication. https://doi.org/10.1016/j.cell.2019.06.012

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. Iii. Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America*, *96*(4), 2076–2087. https://doi.org/10.1121/1.410149

Livosky, M., & Sugar, J. A. (1992). Developmental differences in memory and practice effects in relearning. *Bulletin of the Psychonomic Society*, *30*(3), 205–208. https://doi.org/10.3758/BF03330442

Lukács, Á., & Kemény, F. (2015). Development of different forms of skill learning throughout the lifespan. *Cognitive Science, 39*(2), 383–404. https://doi.org/10.1111/cogs.12143

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), 1–182.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule Learning by Seven-Month-Old Infants. *Science*, *283*(5398), 77. https://doi.org/10.1126/science.283.5398.77

Marmie, W. R., & Healy, A. F. (1995). The Long-Term Retention of a Complex Skill. In A. F. Healy (Ed.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 30–65). SAGE Publications. https://doi.org/10.4135/9781483326887.n2

Mathôt, S. (2017, May 15). Bayes like a Baws: Interpreting Bayesian Repeated Measures in JASP. *Cognitive science and more*. https://www.cogsci.nl/blog/interpreting-bayesian-repeated-measures-in-jasp

McCandliss, B. D., Fiez, J. A., Protopapas, A., Conway, M., & McClelland, J. L.(2002). Success and failure in teaching the r-l contrast to Japanese adults: Tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cognitive, Affective & Behavioral Neuroscience*, *2*(2), 89–108. https://doi.org/10.3758/CABN.2.2.89

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, *102*(3), 419–457.

Melchers, P., & Melchers, M. (2015). *KABC-II: Kaufman Assessment Battery for Children – II. Deutschsprachige Fassung der KABC-II nach A. Kaufman und N.L. Kaufman* (1st edition). Frankfurt: Pearson.

Melchers, P., Schürmann, S., & Scholten, S. (2006). *K-TIM: Kaufman – Test zur Intelligenzmessung für Jugendliche und Erwachsene (K-TIM). Deutschsprachige Fassung des Kaufman – Adolescent and Adult Intelligence Test (KAIT) von A.S. Kaufman und N.L. Kaufman*. Leiden: PITS.

Meulemans, T., van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*(3), 199–221. https://doi.org/10.1006/jecp.1998.2442

Milne, A. E., Petkov, C. I., & Wilson, B. (2018). Auditory and Visual Sequence Learning in Humans and Monkeys using an Artificial Grammar Learning Paradigm. *Neuroscience*, *389*, 104–117. https://doi.org/10.1016/j.neuroscience.2017.06.059

Misyak, J. B., & Christiansen, M. H. (2012). Statistical Learning and Language: An Individual Differences Study. *Language Learning*, *62*(1), 302–331. https://doi.org/10.1111/j.1467-9922.2010.00626.x

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.

Mueller, J. L., Friederici, A. D., & Männel, C. (2018). Developmental changes in automatic rule-learning mechanisms across early childhood. *Developmental Science*, e12700. https://doi.org/10.1111/desc.12700

Murre, J. M. J., & Dros, J. (2015). Replication and Analysis of Ebbinghaus' Forgetting Curve. *PloS One*, *10*(7), e0120644. https://doi.org/10.1371/journal.pone.0120644

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*, 147. https://doi.org/10.3389/fncom.2013.00147

Ngo, C. T., Newcombe, N. S., & Olson, I. R. (2018). The ontogeny of relational memory and pattern separation. *Developmental Science*, *21*(2). https://doi.org/10.1111/desc.12556

Nicolson, R. I., & Fawcett, A. J. (2007). Procedural learning difficulties: Reuniting the developmental disorders? *Trends in Neurosciences*, *30*(4), 135–141. https://doi.org/10.1016/j.tins.2007.02.003

Nicolson, R. I., & Fawcett, A. J. (2011). Dyslexia, dysgraphia, procedural learning and the cerebellum. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *47*(1), 117–127. https://doi.org/10.1016/j.cortex.2009.08.016

Nowak, I., & Baggio, G. (2017). Developmental Constraints on Learning Artificial Grammars with Fixed, Flexible and Free Word Order. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.01816

O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science*, *38*(6), 1229–1248. https://doi.org/10.1111/j.1551-6709.2011.01214.x

Pagel, B., Heed, T., & Röder, B. (2009). Change of reference frame for tactile localization during child development. *Developmental Science*, *12*(6), 929–937. https://doi.org/10.1111/j.1467-7687.2009.00845.x

Palmer, S. D., & Mattys, S. L. (2016). Speech segmentation by statistical learning is supported by domain-general processes within working memory. *Quarterly Journal of Experimental Psychology*, *69*(12), 2390–2401. https://doi.org/10.1080/17470218.2015.1112825

Parkin, A. J., & Streete, S. (1988). Implicit and explicit memory in young children and adults. *British Journal of Psychology*, *79*(3), 361–369.

Pavlidou, E. V., & Bogaerts, L. (2019). Implicit Statistical Learning Across Modalities and Its Relationship With Reading in Childhood. *Frontiers in Psychology*, *10*, 1348. https://doi.org/10.3389/fpsyg.2019.01834

Pavlidou, E. V., & Williams, J. M. (2014). Implicit learning and reading: Insights from typical children and children with developmental dyslexia using the artificial grammar learning (AGL) paradigm. *Research in Developmental Disabilities*, *35*(7), 1457–1472. https://doi.org/10.1016/j.ridd.2014.03.040

Pelz, M., & Kidd, C. (2020). The elaboration of exploratory play. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *375*(1803), 20190503. https://doi.org/10.1098/rstb.2019.0503

Perfors, A. (2012). When do memory limitations lead to regularization? An experimental and computational investigation. *Journal of Memory and Language*, *67*(4), 486–506. https://doi.org/10.1016/j.jml.2012.07.009

Perruchet, P. (2019). What Mechanisms Underlie Implicit Statistical Learning? Transitional Probabilities Versus Chunks in Language Learning. *Topics in Cognitive Science*, *11*(3), 520–535. https://doi.org/10.1111/tops.12403

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

Petermann, F. (2018). *SET 5-10: Sprachstandserhebungstest für Kinder im Alter zwischen 5 und 10 Jahren* (3rd edition). Göttingen: Hogrefe.

Petermann, F. (2012). *WAIS-IV: Wechsler Adult Intelligence Scale - Fourth Edition. Deutschsprachige Adaption der WAIS-IV nach D. Wechsler* (1st edition). Frankfurt: Pearson.

Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia*, *41*(3), 245–251. https://doi.org/10.1016/s0028-3932(02)00157-4

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, *133*(2), 227–244. https://doi.org/10.1037/0033-2909.133.2.227

Qi, Z., Sanchez Araujo, Y., Georgan, W. C., Gabrieli, J. D. E., & Arciuli, J. (2019). Hearing Matters More Than Seeing: A Cross-Modality Study of Statistical Learning and Reading Ability. *Scientific Studies of Reading*, *23*(1), 101–115. https://doi.org/10.1080/10888438.2018.1485680

Quintana, D. S., & Williams, D. R. (2018). Bayesian alternatives for common null-hypothesis significance tests in psychiatry: A non-technical guide using JASP. *BMC Psychiatry*, *18*(1), 178. https://doi.org/10.1186/s12888-018-1761-4

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Sciences*, *11*(7), 274–279. https://doi.org/10.1016/j.tics.2007.05.007

Raviv, L., & Arnon, I. (2017). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality ‐ based differences in the effect of age. *Developmental Science*, e12593. https://doi.org/10.1111/desc.12593

R Core Team (2021). R: A language and environment for statistical computing [Computer Software]. R Foundation for Statistical Computing, Vienna, Austria.URL https://www.R-project.org/

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

Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*(10), 2026–2042. https://doi.org/10.1016/j.neuropsychologia.2013.06.019

Ribordy, F., Jabès, A., Banta Lavenex, P., & Lavenex, P. (2013). Development of allocentric spatial memory abilities in children from 18 months to 5 years of age. *Cognitive Psychology*, *66*(1), 1–29. https://doi.org/10.1016/j.cogpsych.2012.08.001

Rieser, S., & Liepmann, D. (2014). *D-PA: Deutschtest für die Personalauswahl*. Göttingen: Hogrefe.

Röder, B., Heed, T., & Badde, S. (2014). Development of the spatial coding of touch: Ability vs. Automaticity. *Developmental Science*, *17*(6), 944–945. https://doi.org/10.1111/desc.12186

Röder, B., Kekunnaya, R., & Guerreiro, M. J. S. (2021). Neural mechanisms of visual sensitive periods in humans. *Neuroscience and Biobehavioral Reviews*, *120*, 86–99. https://doi.org/10.1016/j.neubiorev.2020.10.030

Rohlf, S., Habets, B., Frieling, M. von, & Röder, B. (2017). Infants are superior in implicit crossmodal learning and use other learning mechanisms than adults. *ELife*, *6*. https://doi.org/10.7554/eLife.28166

Rohlf, S., Li, L., Bruns, P., & Röder, B. (2020). Multisensory Integration Develops Prior to Crossmodal Recalibration. *Current Biology : CB*, *30*(9), 1726-1732.e7. https://doi.org/10.1016/j.cub.2020.02.048

Romano, J. C., Howard, J. H., & Howard, D. V. (2010). One-year retention of general and sequence-specific skills in a probabilistic, serial reaction time task. *Memory (Hove, England)*, *18*(4), 427–441. https://doi.org/10.1080/09658211003742680

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Rosas, R., Ceric, F., Tenorio, M., Mourgues, C., Thibaut, C., Hurtado, E., & Aravena, M. T. (2010). Adhd children outperform normal children in an artificial grammar Implicit learning task: Erp and RT evidence. *Consciousness and Cognition*, *19*(1), 341–351. https://doi.org/10.1016/j.concog.2009.09.006

Saffran, J. R ., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Sameroff, A. J., & Haith, M. M. (Eds.). (1996). *The five to seven year shift: The age of reason and responsibility.* Chicago, IL: University of Chicago Press.

Savion-Lemieux, T., Bailey, J. A., & Penhune, V. B. (2009). Developmental contributions to motor sequence learning. *Experimental Brain Research*, *195*(2), 293–306. https://doi.org/10.1007/s00221-009-1786-5

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning.

Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, *372*(1711). https://doi.org/10.1098/rstb.2016.0049

Scherag, A., Demuth, L., Rösler, F., Neville, H. J., & Röder, B. (2004). The effects of late acquisition of L2 and the consequences of immigration on L1 for semantic and morpho-syntactic language aspects. *Cognition*, *93*(3), B97-108. https://doi.org/10.1016/j.cognition.2004.02.003

Schiff, R., & Katan, P. (2014). Does complexity matter? Meta-analysis of learner performance in artificial grammar tasks. *Frontiers in Psychology*, *5*, 1084. https://doi.org/10.3389/fpsyg.2014.01084

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal Structure Predicts Statistical Learning and Associative Inference Abilities during Development. *Journal of Cognitive Neuroscience*, *29*(1), 37–51. https://doi.org/10.1162/jocn_a_01028

Schommartz, I., Lembcke, P. F., Pupillo, F., Schuetz, H., Chamorro, N. W. de, Bauer, M., Kaindl, A. M., Buss, C., & Shing, Y. L. (2023). Distinct multivariate structural brain profiles are related to variations in short- and long-delay memory consolidation across children and young adults. *Developmental Cognitive Neuroscience*, *59*, 101192. https://doi.org/10.1016/j.dcn.2022.101192

Schönbrodt, F. D., & Wagenmakers, E.-J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin & Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Scimeca, J. M., & Badre, D. (2012). Striatal contributions to declarative memory retrieval. *Neuron*, *75*(3), 380–392. https://doi.org/10.1016/j.neuron.2012.07.014

Shing, Y. L., Finke, C., Hoffmann, M., Pajkert, A., Heekeren, H. R., & Ploner, C. J. (2019). Integrating across memory episodes: Developmental trends. *PloS One*, *14*(4), e0215848. https://doi.org/10.1371/journal.pone.0215848

Shufaniya, A., & Arnon, I. (2018). Statistical Learning Is Not Age-Invariant During Childhood: Performance Improves With Age Across Modality. *Cognitive Science*, *42*(8), 3100–3115. https://doi.org/10.1111/cogs.12692

Siegelman, N., Bogaerts, L., Armstrong, B. C., & Frost, R. (2019). What exactly is learned in visual statistical learning? Insights from Bayesian modeling. *Cognition*, *192*, 104002. https://doi.org/10.1016/j.cognition.2019.06.014

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition, 177*, 198–213. https://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: Current pitfalls and possible solutions. *Behavior Research Methods*, *49*(2), 418–432. https://doi.org/10.3758/s13428-016-0719-z

Skosnik, P. D., Mirza, F., Gitelman, D. R., Parrish, T. B., Mesulam, M.-M., & Reber, P. J. (2002). Neural correlates of artificial grammar learning. *NeuroImage, 17*(3), 1306–1314.

Smalle, E. H. M., Daikoku, T., Szmalec, A., Duyck, W., & Möttönen, R. (2022). Unlocking adults' implicit statistical learning by cognitive depletion. *Proceedings of the National Academy of Sciences of the United States of America, 119*(2). https://doi.org/10.1073/pnas.2026011119

Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M., & Szmalec, A. (2017). Children retain implicitly learned phonological sequences better than adults: A longitudinal study. *Developmental Science, 21*(5), e12634. https://doi.org/10.1111/desc.12634

Smalle, E. H. M., Panouilleres, M., Szmalec, A., & Möttönen, R. (2017). Language learning in the adult brain: Disrupting the dorsolateral prefrontal cortex facilitates word-form learning. *Scientific Reports*, *7*(1), 13966. https://doi.org/10.1038/s41598-017-14547-x

Smith, A. C., Stefani, M. R., Moghaddam, B., & Brown, E. N. (2005). Analysis and design of behavioral experiments to characterize population learning. *Journal of Neurophysiology*, *93*(3), 1776–1792. https://doi.org/10.1152/jn.00765.2004

Tandoc, M., Nadendla, B., Pham, T., & Finn, A. (2022). *Dividing attention hurts learning in adults but not children. Advance online publication.* PsyArXiv. https://doi.org/10.31234/osf.io/aqw28

Taylor, J. E., Cortese, A., Barron, H. C., Pan, X., Sakagami, M., & Zeithamova, D. (2021). How do we generalize? *Neurons, Behavior, Data Analysis and Theory*, *1*. https://doi.org/10.51628/001c.27687

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

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Worden, M. S. (2004). Evidence of Developmental Differences in Implicit Sequence Learning: An fMRI Study of Children and Adults. *Journal of Cognitive Neuroscience*, *16*(8), 1339–1351. https://doi.org/10.1162/0898929042304688

Torkildsen, J. v. K., Arciuli, J., & Wie, O. B. (2019). Individual differences in statistical learning predict children's reading ability in a semi-transparent orthography. *Learning and Individual Differences*, *69*, 60–68. https://doi.org/10.1016/j.lindif.2018.11.003

Tóth-Fáber, E., Janacsek, K., & Németh, D. (2021). Statistical and sequence learning lead to persistent memory in children after a one-year offline period. *Scientific Reports*, *11*(1), 12418. https://doi.org/10.1038/s41598-021-90560-5

Tóth-Fáber, E., Nemeth, D., & Janacsek, K. (2023). Lifespan developmental invariance in memory consolidation: Evidence from procedural memory. *PNAS Nexus*, *2*(3), pgad037. https://doi.org/10.1093/pnasnexus/pgad037

Tóth-Fáber, E., Tárnok, Z., Takács, Á., Janacsek, K., & Németh, D. (2021). Access to Procedural Memories After One Year: Evidence for Robust Memory Consolidation in Tourette Syndrome. *Frontiers in Human Neuroscience*, *15*, 715254. https://doi.org/10.3389/fnhum.2021.715254

Uddén, J., Folia, V., Forkstam, C., Ingvar, M., Fernandez, G., Overeem, S., van Elswijk, G., Hagoort, P., & Petersson, K. M. (2008). The inferior frontal cortex in artificial syntax processing: An rTMS study. *Brain Research*, *1224*, 69–78. https://doi.org/10.1016/j.brainres.2008.05.070

Uddén, J., Ingvar, M., Hagoort, P., & Petersson, K. M. (2017). Broca’s region: A causal role in implicit processing of grammars with crossed non-adjacent dependencies. *Cognition*, *164*, 188–198. https://doi.org/10.1016/j.cognition.2017.03.010

Uddén, J., & Männel, C. (2018). *Artificial Grammar Learning and Its Neurobiology in Relation to Language Processing and Development - Oxford Handbooks* (Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.33

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews. Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1-2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *41*(3), 399–433.

van der Kant, A., Männel, C., Paul, M., Friederici, A. D., Höhle, B., & Wartenburger, I. (2020). Linguistic and non-linguistic non-adjacent dependency learning in early development. *Developmental Cognitive Neuroscience*, *45*, 100819. https://doi.org/10.1016/j.dcn.2020.100819

van Witteloostuijn, M., Boersma, P., Wijnen, F., & Rispens, J. (2017). Visual artificial grammar learning in dyslexia: A meta-analysis. *Research in Developmental Disabilities*, *70*, 126–137. https://doi.org/10.1016/j.ridd.2017.09.006

Visser, I., Raijmakers, M. E. J., & Pothos, E. M. (2009). Individual strategies in artificial grammar learning. *The American Journal of Psychology*, *122*(3), 293–307.

Vlach, H. A. (2014). The Spacing Effect in Children's Generalization of Knowledge: Allowing Children Time to Forget Promotes Their Ability to Learn. *Child Development Perspectives*, *8*(3), 163–168. https://doi.org/10.1111/cdep.12079

Vries, M. H. de, Barth, A. C. R., Maiworm, S., Knecht, S., Zwitserlood, P., & Floel, A. (2010). Electrical stimulation of Broca's area enhances implicit learning of an artificial grammar. *Journal of Cognitive Neuroscience*, *22*(11), 2427–2436. https://doi.org/10.1162/jocn.2009.21385

Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *The Behavioral and Brain Sciences*, *28*(1), 51-64; discussion 64-104. https://doi.org/10.1017/s0140525x05000026

Weiermann, B., & Meier, B. (2012). Incidental sequence learning across the lifespan. *Cognition*, *123*(3), 380–391. https://doi.org/10.1016/j.cognition.2012.02.010

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*, 173–196. https://doi.org/10.1146/annurev-psych-010814-015104

West, G., Melby-Lervåg, M., & Hulme, C. (2021). Is a procedural learning deficit a causal risk factor for developmental language disorder or dyslexia? A meta-analytic review. *Developmental Psychology*, *57*(5), 749–770. https://doi.org/10.1037/dev0001172

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: We see some problems. *Developmental Science*, *21*(2). https://doi.org/10.1111/desc.12552

West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2019). The procedural deficit hypothesis of language learning disorders: We still see some serious problems. *Developmental Science*, *22*(4), e12813. https://doi.org/10.1111/desc.12813

White, E. J., Hutka, S. A., Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: Implications for language acquisition, music training and transfer across the lifespan. *Frontiers in Systems Neuroscience*, *7*, 90. https://doi.org/10.3389/fnsys.2013.00090

Whitmarsh, S., Uddén, J., Barendregt, H., & Petersson, K. M. (2013). Mindfulness reduces habitual responding based on implicit knowledge: Evidence from artificial grammar learning. *Consciousness and Cognition*, *22*(3), 833–845. https://doi.org/10.1016/j.concog.2013.05.007

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation – What can be learnt from children? *Neuroscience & Biobehavioral Reviews*, *36*(7), 1718–1728. https://doi.org/10.1016/j.neubiorev.2012.03.002

Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., & Born, J. (2013). The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge. *Nature Neuroscience*, *16*(4), 391–393. https://doi.org/10.1038/nn.3343

Wilson, B., Spierings, M., Ravignani, A., Mueller, J. L., Mintz, T. H., Wijnen, F., van der Kant, A., Smith, K., & Rey, A. (2020). Non-adjacent Dependency Learning in Humans and Other Animals. *Topics in Cognitive Science*, *12*(3), 843–858. https://doi.org/10.1111/tops.12381

Witt, A., Puspitawati, I., & Vinter, A. (2013). How explicit and implicit test instructions in an implicit learning task affect performance. *PloS One*, *8*(1), e53296. https://doi.org/10.1371/journal.pone.0053296

Witt, A., & Vinter, A. (2012). Artificial grammar learning in children: Abstraction of rules or sensitivity to perceptual features? *Psychological Research*, *76*(1), 97–110. https://doi.org/10.1007/s00426-011-0328-5

Witt, A., & Vinter, A. (2017). Perceptual and Positional Saliencies Influence Children's Sequence Learning Differently with Age and Instructions at Test. *Quarterly Journal of Experimental Psychology*, *70*(11), 2219–2233. https://doi.org/10.1080/17470218.2016.1230141

Wonnacott, E. (2011). Balancing generalization and lexical conservatism: An artificial language study with child learners. *Journal of Memory and Language*, *65*(1), 1–14. https://doi.org/10.1016/j.jml.2011.03.001

Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., Jones, T., & Zuo, Y. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, *462*(7275), 915–919. https://doi.org/10.1038/nature08389

Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature, 462*(7275), 920–924. https://doi.org/10.1038/nature08577

Yassa, M. A., & Stark, C. E. L. (2011). Pattern separation in the hippocampus. *Trends in Neurosciences*, *34*(10), 515–525. https://doi.org/10.1016/j.tins.2011.06.006

Yasumatsu, N., Matsuzaki, M., Miyazaki, T., Noguchi, J., & Kasai, H. (2008). Principles of long-term dynamics of dendritic spines. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *28*(50), 13592–13608. https://doi.org/10.1523/JNEUROSCI.0603-08.2008

Zinszer, B. D., Hannon, J., Kouadio, A. É., Akpé, H., Tanoh, F., Hu, A., Qi, Z., & Jasińska, K. K. (2020). *Statistical learning in children's emergent L2 literacy: Cross-cultural insights from rural Côte d'Ivoire.* AfricArcXiv. https://doi.org/10.31730/osf.io/q8k5w

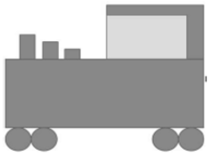
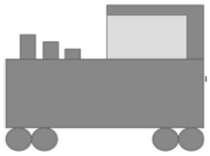
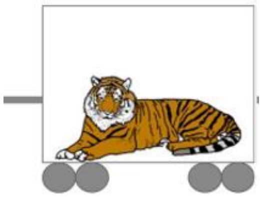
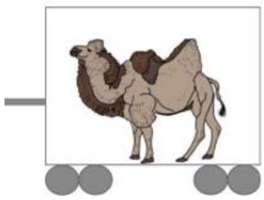
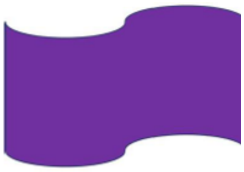
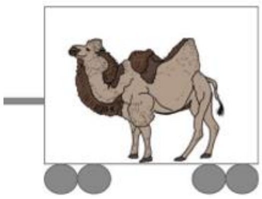
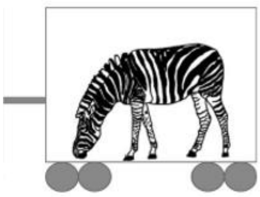
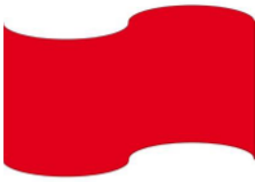
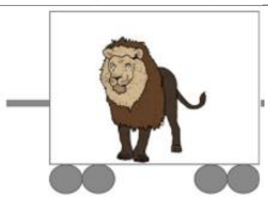
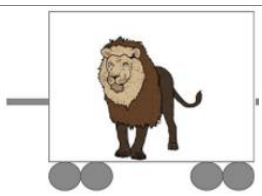
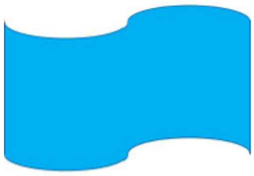
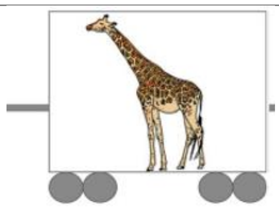
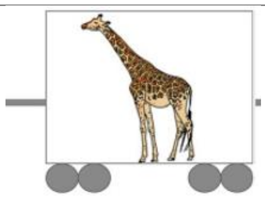
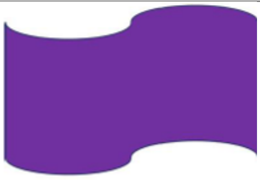
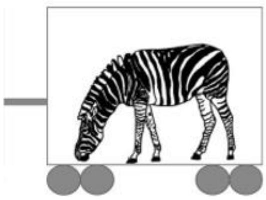
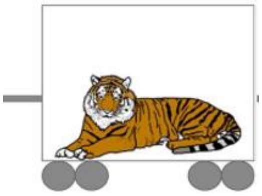
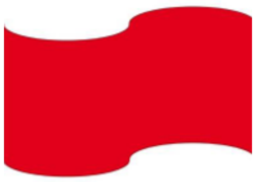
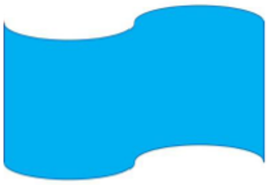
Contents Appendix

APPENDIX A:
Stimuli, Instructions & Questionnaires

Stimulus Sets for AGL Task

Table A.1

Picture Assignment in AGL Task, corresponding to Numbers in Figure 2

# in Fig. 2	Assigned picture Stimulus Set Animals, Version 1	Assigned picture Stimulus Set Animals, Version 2	Assigned picture Stimulus Set Colors, Version 1	Assigned picture Stimulus Set Colors, Version 2
1				
2				
3				
4				
5				
6				

Note. Picture assignments for both task versions per stimulus set were randomly picked from all possible assignments from permutations generated by a randomization script. The resulting task versions were applied in a counterbalanced manner in all age groups (described in detail in the *Methods* section of Chapter II).

AGL Task: Grammatical & Ungrammatical Sequences

Table A.2

Sequences presented in the AGL task

3-5 item grammatical sequences	**3-5 item ungrammatical sequences**	**6-7 item grammatical sequences**	**6-7 item ungrammatical sequences**
224	242	566534	256624
554	2432	233324	225664
5534	2423	553654	555364
5654	2342	566654	262534
2324	5435	232654	565634
22654	56354	226654	2332564
56534	25624	226534	2263354
23324	22564	2333324	2362534
56654	53564	2326534	2623354
	26254	5666534	2636524
	55364	5666654	2353264
		5536654	2566234
		5536534	5355664
		2326654	5566354
		2266534	2326354
		2266654	2236564
		5653654	2652364
		2332654	5535664
			5536354

Note. Numbers were replaced by pictures of animals/color segments, see stimuli Table A.1. Grammatical sequences followed AG rules (see Fig. 2) and were presented in learning and test phases. Ungrammatical sequences violated AG rules and were presented in test phases (for details, see *Methods* section in Chapter II).

Instructions for AGL Task I: Stimulus Set Animals

Lernphase 1:

„Gleich siehst du den Zirkus von Frau Pepe. Sie reist mit dem Zirkus in einem Zug durch das ganze Land. In jedem Zugwaggon ist ein Zirkustier. Frau Pepe stellt den Zug so zusammen, dass sich die Tiere miteinander wohlfühlen. Du bist ein Detektiv und wirst nun einige Züge von Frau Pepe sehen. Schau sie dir genau an, wir stellen später Fragen dazu."

Testphase 1:

„Du kommst als Detektiv zum Bahnhof und siehst dort zwei Zirkuszüge. Nur einer der Züge gehört Frau Pepe. Hilf uns herauszufinden, welcher der beiden von ihr zusammengestellt wurde. Gehe nach deinem Bauchgefühl und wähle den Zug aus, der dir als erstes in den Sinn kommt.
Um den Zug auszuwählen, tippe ihn einfach auf dem Bildschirm an."

Bei jedem Testtrial (2 Züge) wird wiederholt:

„Welcher Zirkuszug wurde von Frau Pepe zusammengestellt? Tippe ihn auf dem Bildschirm an."

Lernphase 2-5:

„Der Zirkus von Frau Pepe reist nun weiter. Frau Pepe stellt den Zug immer noch auf dieselbe Art und Weise zusammen. Nämlich so, dass sich die Tiere miteinander wohlfühlen. Du wirst nun nochmal einige Züge von Frau Pepe sehen. Schau sie dir genau an, wir stellen später nochmal Fragen dazu."

Testphase 2-5:

„Du kommst wieder als Detektiv zum Bahnhof und siehst dort zwei Zirkuszüge. Nur einer der Züge gehört Frau Pepe. Hilf uns herauszufinden, welcher der beiden von ihr zusammengestellt wurde. Gehe nach deinem Bauchgefühl und wähle den Zug aus, der dir als erstes in den Sinn kommt.
Um den Zug auszuwählen, tippe ihn einfach auf dem Bildschirm an."

Bei jedem Testtrial (2 Züge) wird wiederholt:

„Welcher Zirkuszug wurde von Frau Pepe zusammengestellt? Tippe ihn auf dem Bildschirm an."

Instructions for AGL Task II: Stimulus Set Colors

Lernphase 1:

„Gleich siehst du die Fahnen der Sportmannschaft „Starke Tiger“. Die Fahnen bestehen aus bunten Farben und werden auf Sportturnieren aufgestellt. Auf jedes Turnier bringen die „Starken Tiger“ eine andere Fahne mit, damit ihren Fans nicht langweilig wird. Ihre Fahnen stellen die „Starken Tiger“ so zusammen, dass die Farben gut zueinander passen. Du bist ein Detektiv und wirst nun einige Fahnen der „Starken Tiger“ sehen. Schau sie dir genau an, wir stellen später Fragen dazu.“

Testphase 1:

„Du kommst als Detektiv zu einem Turnier und siehst dort zwei Fahnen. Nur eine davon gehört der Sportmannschaft „Starke Tiger“. Hilf uns herauszufinden, welche der beiden von ihnen zusammengestellt wurde. Gehe nach deinem Bauchgefühl und wähle die Fahne aus, die dir als erstes in den Sinn kommt.
Um die Fahne auszuwählen, tippe sie einfach auf dem Bildschirm an.“

Bei jedem Testtrial (2 Züge) wird wiederholt:

„Welche Fahne wurde von den „Starken Tigern“ zusammengestellt? Tippe sie auf dem Bildschirm an.“

Lernphase 2-5:

„Die „Starken Tiger“ bringen ihre Fahnen auf weitere Turniere mit. Sie stellen ihr Fahnen immer noch auf dieselbe Art und Weise zusammen. Nämlich so, dass die Farben gut zueinander passen. Du wirst nun nochmal einige Fahnen der „Starken Tiger“ sehen. Schau sie dir genau an, wir stellen später nochmal Fragen dazu.“

Testphase 2-5:

„Du kommst wieder als Detektiv zu einem Turnier und siehst dort zwei Fahnen. Nur eine davon gehört der Sportmannschaft „Starke Tiger“. Hilf uns herauszufinden, welche der beiden von ihnen zusammengestellt wurde. Gehe nach deinem Bauchgefühl und wähle die Fahne aus, die dir als erstes in den Sinn kommt.
Um die Fahne auszuwählen, tippe sie einfach auf dem Bildschirm an.“

Bei jedem Testtrial (2 Züge) wird wiederholt:

„Welche Fahne wurde von den „Starken Tigern“ zusammengestellt? Tippe sie auf dem Bildschirm an.“

Questionnaire to assess explicit sequence knowledge: one example

(for Stimulus Set Animals, assessed in adults – adapted versions were applied for Stimulus Set Colors, and shortened versions for children in both stimulus sets)

<u>Fragebogen explizites Wissen (Grammatik: Reber 1967) – Tiere 1 + 2</u>

angepasst von: Whitmarsh, S., Uddén, J., Barendregt, H. & Petersson, K. M. (2013). Mindfulness reduces habitual responding based on implicit knowledge. Evidence from artificial grammar learning. *Consciousness and cognition, 22* (3), 833-845.

1) Allgemeine Fragen
   - *„Was denkst du, worum es in diesem Spiel ging?“*

     ____

     ____

   - *„Hattest du eine Strategie, um zu entscheiden welcher Zug Frau Pepe gehört?“*

     ____

     ____

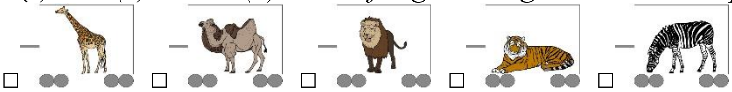
   - *„Ist dir etwas an den Zirkuszügen von Frau Pepe aufgefallen?“*

     ____

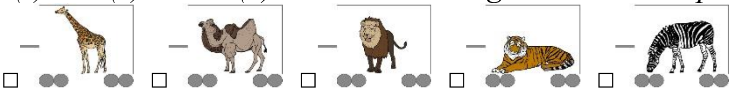
     ____

Questionnaire to assess explicit sequence knowledge: one example *(continued)*

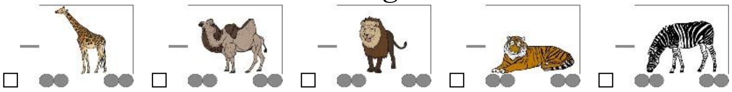
2) Spezifische Fragen
   a) **„** *Welche(s) Tier(e) konnte(n) am Anfang des Zuges von Frau Pepe sein?“*

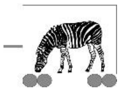


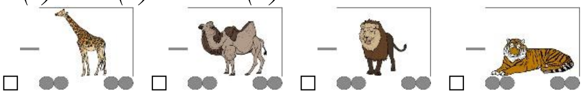
   b) *„Welche(s) Tier(e) konnte(n) am Ende des Zuges von Frau Pepe sein?“*

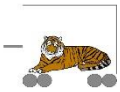


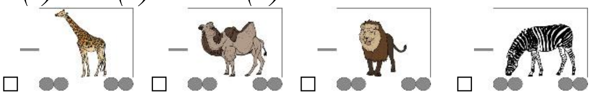
   c) *„Von welchem Tier konnten mehrere gleiche hintereinander kommen?“*

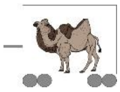


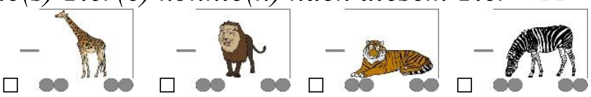
   d) *„Welche(s) Tier(e) konnte(n) nach diesem Tier*  *kommen?“*

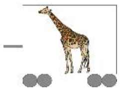


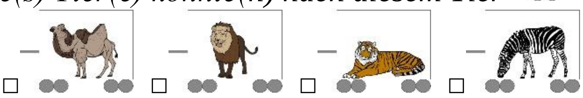
   e) *„Welche(s) Tier(e) konnte(n) nach diesem Tier*  *kommen?“*



   f) *„Welche(s) Tier(e) konnte(n) nach diesem Tier*  *kommen?“*

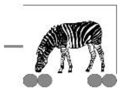


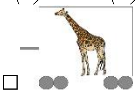
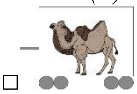
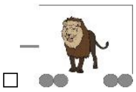
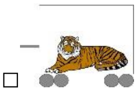
   g) *„Welche(s) Tier(e) konnte(n) nach diesem Tier*  *kommen?“*

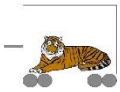


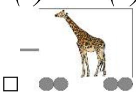
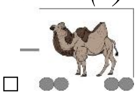
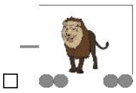
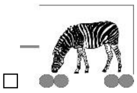
Bitte wenden →

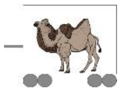
Questionnaire to assess explicit sequence knowledge: one example *(continued)*

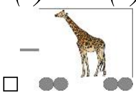
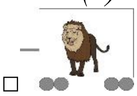
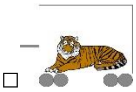
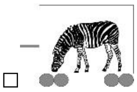
h) *„Welche(s) Tier(e) konnte(n) NICHT nach diesem Tier*  *kommen?“*

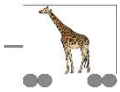
☐  ☐  ☐  ☐ 

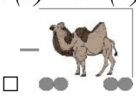
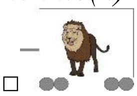
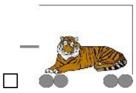
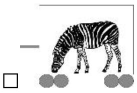
i) *„Welche(s) Tier(e) konnte(n) NICHT nach diesem Tier*  *kommen?“*

☐  ☐  ☐  ☐ 

j) *„Welche(s) Tier(e) konnte(n) NICHT nach diesem Tier*  *kommen?“*

☐  ☐  ☐  ☐ 

k) *„Welche(s) Tier(e) konnte(n) NICHT nach diesem Tier*  *kommen?“*

☐  ☐  ☐  ☐ 

APPENDIX B:
Additional Analyses Chapter II (Project 1)

1. **Descriptive data for control analyses without the first AGL task block per session**

Table B.1

A *Year 1: Proportion Correct Across Block 2-5 in AGL per Session and Age Group*

	Session 1		**Session 2**		**Session 3**		**Transfer 1**	
	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1
N	27	28	27	28	27	28	27	28
M	.56	.82	.62	.91	.63	.94	.64	.90
SD	.11	.10	.12	.07	.17	.07	.11	.11
Min	.23	.63	.40	.78	.37	.76	.33	.65
Max	.73	1.00	.85	1.00	.95	1.00	.88	1.00
*N**	16	20	16	20	16	20	16	20
M	.58	.81	.67	.93	.68	.95	.68	.91
SD	.08	.11	.12	.06	.18	.07	.09	.12
Min	.45	.63	.50	.80	.43	.78	.53	.65
Max	.73	.98	.85	1.00	.95	1.00	.88	1.00

B *Year 2: Proportion Correct in AGL Across Block 2-5 per Session and Age Group*

	Session 4		**Session 5**		**Session 6**		**Transfer 2**	
	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1	7yo	Ad 1
N	16	20	16	20	16	20	16	20
M	.73	.93	.76	.94	.79	.94	.73	.98
SD	.11	.10	.17	.10	.15	.09	.14	.09
Min	.53	.60	.43	.68	.48	.70	.48	.70
Max	.93	1.00	1.00	1.00	1.00	1.00	.98	1.00

Note. 7yo = 7-year-olds, Ad 1 = Adults 1, Min = minimal value, Max = maximal value

* Subgroup of returning participants with data for Year 1 & Year 2, see *Participants* in Chapter II.

2. AGL performance associations with levels of explicit knowledge in Year 1

Explicit knowledge about sequence rules (assessed at the end of Transfer 1) was marginally positively correlated with a performance increase in the transfer session compared to initial performance with the first stimulus set (Transfer Savings: Transfer 1 – Session 1) in 7-year-olds (r_s = .42, p = .058, BF_{10} = 1.68). For adults, this correlation did not reach the significance level (r_s = .25; p = .193, BF_{10} = .55). The second transfer difference (Transfer 1 – Session 3) as preserved performance from the well-practiced first stimulus set did not correlate significantly with explicit sequence knowledge in either age group (7-year-olds: r_s = .07, p = .725, BF_{10} = .26; adults: r_s = .24, p = .448, BF_{10} = .42). For the performance increase from Session 1 to Session 3, children showed a numerically larger positive correlation with explicit sequence knowledge (r_s = .33; p = .176, BF_{10} = 1.11) than adults (r_s = .10; p = .619, BF_{10} = .31), but none of these correlations was statistically significant. All correlation patterns remained the same when one child with very poor explicit sequence knowledge was excluded (Z-Score < -3; Transfer 1 – Session 1: r_s = .44, p = .075, BF_{10} = 2.15; Transfer 1 – Session 3: r_s = .19, p = .382, BF_{10} = .36; Session 3 – Session 1: r_s = .27, p = .382, BF_{10} = .37).

3. AGL performance associations with levels of explicit knowledge in Year 2

We asked whether explicit sequence knowledge in the end of Year 1 predicts performance increases in Year 2: No significant correlations between explicit knowledge scores at the end of Transfer 1 with an increase over Session 4 to 6 (7-year-olds: r_s = -.22, p = .826, BF_{10} = .56; adults: r_s = -.19, p = .826, BF_{10} = .38), nor with an improvement in Transfer 2 compared to Session 4 (7-year-olds: r_s = .09, p = .728, BF_{10} = .34; adults: r_s = -.38, p = .180, BF_{10} = 1.07), nor with better performance in Transfer 2 compared to Session 6 (7-year-olds: r_s = .27, p = .384, BF_{10} = .85; adults: r_s = -.30, p = .384, BF_{10} = .77) emerged. Correlating explicit knowledge scores from Year 1 with higher start levels (Session 4 – Session 1, Session 4 – Session 3) or higher end levels (Session 6 – Session 3) at relearning in Year 2 compared to initial learning in Year 1, produced no significant associations as well (7-year-olds: all $|r_s| \leq .31$, all $p \geq .472$, $BF_{10} \leq .68$; adults: all $|r_s| \leq .15$, all $p \geq .862$, $BF_{10} \leq .38$).

4. Effects of trial difficulty and short-term familiarity on AGL performance in Year 2

We evaluated for relearning across all four sessions of Year 2, how the two task characteristics (1) short-term familiarity (grammatical sequence of a test trial seen vs. not seen in the preceding learning phase) and (2) difficulty (short test trials with low ACS = simple vs. long test trials with high ACS = difficult, both see *Construction of Grammatical and Ungrammatical Sequences*) relate to task performance in both age groups.

An ANOVA on the influence of short-term familiarity with the factors *Age* (between-subject; levels: 7-year-olds, adults) and *Trial Type* (within-subject; levels: previously seen vs. not seen grammatical sequences in test trials) revealed only a main effect of *Age* ($F(1, 34) = 25.88$, $p < .001$, $\eta^2_{g} = .42$, $BF_{incl} > 100$; both other $F(1, 34) \leq 0.62$, $p \geq .436$, $\eta^2_{g} < .01$, $BF_{incl} \leq .36$), with adults outperforming children in both trial types.

A different result pattern emerged for the ANOVA on trial difficulty with the factors *Age* (between-subject; levels: 7-year-olds, adults) and *Trial Type* (within-subject; levels: simple vs. difficult test trial): This analysis revealed a significant *Age*Trial Type* interaction ($F(1, 34) = 9.98$, $p = .003$, $\eta^2_{g} = .04$, $BF_{incl} = .84$), in addition to main effects of *Age* and *Trial Type* (both $F(1, 34) \geq 25.99$, $p < .001$, $\eta^2_{g} \geq .13$, $BF_{incl} > 100$): Adults outperformed children in both trial types and participants performed poorer in long test trials with more shared transitions between grammatical and ungrammatical sequences (main effects of *Age* and *Trial Type*). Additionally, this effect of trial difficulty was less pronounced in adults than in 7-year-olds ($V = 247.5$, $p = .006$, $r = .54$, $BF_{10} = 3.11$).

Despite this effect of trial difficulty, also the group of 7-year-olds was able to learn the more challenging type of sequences in Year 2 (long test trials with high ACS: $t(15) = 4.75$, $p < .001$, $d = 1.19$, $BF_{10} > 100$).

APPENDIX C: Additional Analyses Chapter III (Project 2)

1. **Descriptive data for control analyses without the first AGL task block per session**

Table C.1

A *Year 1: Proportion Correct Across Block 2-5 in AGL per Session and Age Group*

	Session 1			**Session 2**			**Session 3**		
	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2
N	24	27	20	24	27	20	24	27	20
M	.51	.51	.79	.51	.57	.88	.52	.65	.91
SD	.08	.11	.13	.09	.12	.13	.10	.11	.12
Min	.38	.38	.58	.38	.30	.60	.35	.38	.60
Max	.68	.85	.95	.70	.75	1.00	.78	.95	1.00

B *Year 2: Proportion Correct Across Block 2-5 in AGL per Session and Age Group*

	Session 4			**Session 5**			**Session 6**			**Transfer 2**		
	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2	5yo	6yo	Ad 2
N	24	27	20	24	27	20	24	27	20	24	27	20
M	.55	.62	.88	.61	.65	.93	.64	.72	.92	.60	.63	.89
SD	.11	.13	.11	.11	.15	.08	.12	.14	.11	.10	.12	.14
Min	.43	.35	.63	.43	.40	.75	.48	.48	.63	.45	.36	.53
Max	.83	.93	1.00	.80	.95	1.00	1.00	1.00	1.00	.75	.80	1.00

Note. 5yo = 5-year-olds, 6yo = 6-year-olds, Ad 2 = Adults 2, Min = minimal value, Max = maximal value

2. **Effects of trial difficulty and short-term familiarity on AGL performance in Year 1**

We evaluated for learning in all three sessions of Year 1 taken together, how the two AGL task characteristics (1) short-term familiarity (grammatical sequence of a test trial seen vs. not seen in the preceding learning phase) and (2) difficulty (short test trials with low ACS = easy vs. long test trials with high ACS = difficult, both see *Construction of Grammatical and Ungrammatical Sequences* in Chapter II) relate to task performance in the three age groups.

An ANOVA on the influence of short-term familiarity with the factors *Age* (between-subject; levels: 5-year-olds, 6-year-olds, Adults 2) and *Trial Type* (within-subject; levels:

previously seen vs. not seen grammatical sequences in test trials) revealed only a main effect of *Age* (F(2, 68) = 91.48, p < .001, η^2_{g} = .68; both other $F \leq 1.45$, $p \geq .241$, $\eta^2_{g} \leq .01$). Thus, age groups differed in their overall performance levels, but not in their performance for trials with grammatical sequences that were seen vs. such that were not seen in the directly preceding learning phase.

An ANOVA on the influence of trial difficulty with the factors *Age* (between-subject; levels: 5-year-olds, 6-year-olds, Adults 2) and *Trial Type* (within-subject; levels: easy vs. difficult test trial) revealed main effects of *Age* (F(2, 68) = 96.31, p < .001, η^2_{g} = .68) and of *Trial Type* (F(1, 68) = 13.27, p < .001, η^2_{g} = .05), but no significant *Age*Trial Type* interaction (F(2, 68) = 0.45, p = .641, η^2_{g} < .01). This means that overall performance levels differed depending on age, and that all age groups performed better in easy vs. difficult trials in Year 1. But age groups did not differ in the size of this effect of trial type on performance.

3. **Effects of trial difficulty and short-term familiarity on AGL performance in Year 2**

For relearning in all four sessions of Year 2 taken together, task performance was compared for the two task characteristics (1) short-term familiarity (grammatical sequence of a test trial seen vs. not seen in the preceding learning phase) and (2) difficulty (short test trials with low ACS = easy vs. long test trials with high ACS = difficult, both see *Construction of Grammatical and Ungrammatical Sequences* in Chapter II) in all three age groups.

An ANOVA on the influence of short-term familiarity with the factors *Age* (5-year-olds, 6-year-olds, Adults 2) and *Trial Type* (within-subject; levels: previously seen vs. not seen grammatical sequences in test trials) revealed main effects of *Age* (F(2, 68) = 65.21, p < .001, η^2_{g} = .63) and *Trial Type* (F(1, 68) = 6.45, p = .013, η^2_{g} = .01), but no *Age*Trial Type* interaction (F(2, 68) = 0.60, p = .551, η^2_{g} < .01). This means that overall performance levels differed depending on age in Year 2, and that all age groups performed better in trials with grammatical sequences that they had seen in the preceding learning phase. But age groups did not differ in in the size of this effect of trial type on performance.

An ANOVA on the influence of trial difficulty with the factors *Age* (between-subject; levels: 5-year-olds, 6-year-olds, Adults 2) and *Trial Type* (within-subject; levels: easy vs. difficult test trial) revealed main effects of *Age* (F(2, 68) = 64.66, p < .001, η^2_{g} = .60) and of *Trial Type* (F(1, 68) = 48.98, p < .001, η^2_{g} = .12), but no significant *Age*Trial Type* interaction (F(2, 68) = 1.08, p = .345, η^2_{g} = .01). This means that age groups differed in their overall performance levels, and that all of them performed better in easy vs. difficult trials in

Year 2. But age groups did not differ in the way how trial difficulty affected their performance.

To sum up, trial properties (short-term familiarity & trial difficulty) affected overall AGL task performance in both years, with all age groups performing worse in more challenging trial types (long test trials with high ACS in both years & grammatical sequence of a test trial which had not been seen in the preceding learning phase in Year 2). However, having seen the grammatical sequence of a test trial in the previous learning phase did not matter for performance in the sessions of Year 1. In both years, 5-year-olds, 6-year-olds and Adults 2 did not differ in any of the reported effects of trial types on overall performance.

4. Comparison of Adults 1 and Adults 2 in their transfer of visual regularities

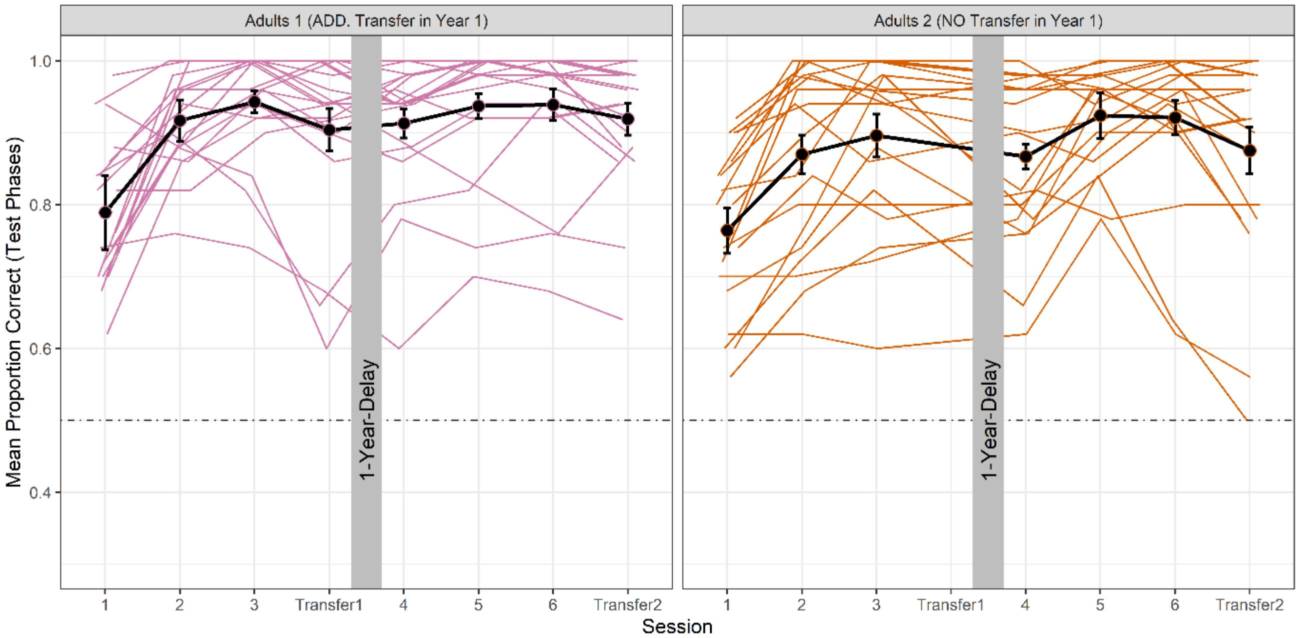
The two adult groups from Project 1 (Adults 1, Chapter II) and from Project 2 (Adults 2, Chapter III) were compared in their transfer to stimulus set 2, to evaluate if the timing of learning sessions with stimulus set 1 mattered for transfer effects. Adults 1 completed two transfer sessions, one in the end of each year, while Adults 2 completed a single transfer session in the end of Year 2 (see Table 1). Learning trajectories for both groups in all completed sessions are depicted in Figure C.1.

To test for group differences in transfer, absolute performance levels in Transfer 2 (Year 2) and in the first transfer session of each group (Adults 1: Transfer 1 in Year 1, Adults 2: Transfer 2 in Year 2) were compared between the two adult groups. Additionally, four transfer effects were calculated as sessions differences that correspond to the transfer measures of Project 1 and Project 2 (Chapter II & III, Transfer Savings: Transfer 2 – Session 4, Transfer Loss: Transfer 2 – Session 6, Transfer 2 – Session 1, first transfer session – Session 1; see Fig. C.2).

Both groups did not differ significantly in their absolute performance levels at transfer, neither at Transfer 2 (*M(SD)* Adults 1: .92(.09), *M(SD)* Adults 2: .88(.14), Welch's $t(32.76) = 1.16$, $p = .253$, $d = .37$, $BF_{10} = .53$), nor in their first transfer session (*M(SD)* Adults 1: .90(.12), *M(SD)* Adults 2: .88(.14), Welch's $t(36.86) = 0.70$, $p = .486$, $d = .22$, $BF_{10} = .38$). Similarly, no statistically significant difference between Adults 1 and Adults 2 emerged from the rmANOVAs which compared Transfer Savings, Transfer Loss, Transfer 2 vs. Session 1, and the first transfer session vs. Session 1 between groups (see Table C.2).

Figure C.1

Performance Trajectories Across Sessions for Adults 1 (left) and Adults 2 (right)



Note. Mean proportion of correct responses in the test phases of each session for adult groups. Learning curves of single participants are depicted in color. The dotted horizontal lines mark chance level performance. Error bars indicate 95% CIs corrected for within-subject comparison according to Morey (2008). ADD. = additional.

Table C.2

Results of rmANOVAs on Transfer Effects with *Group* (Adults 1, Adults 2) as between subject factor & *Session* as within-subject factor

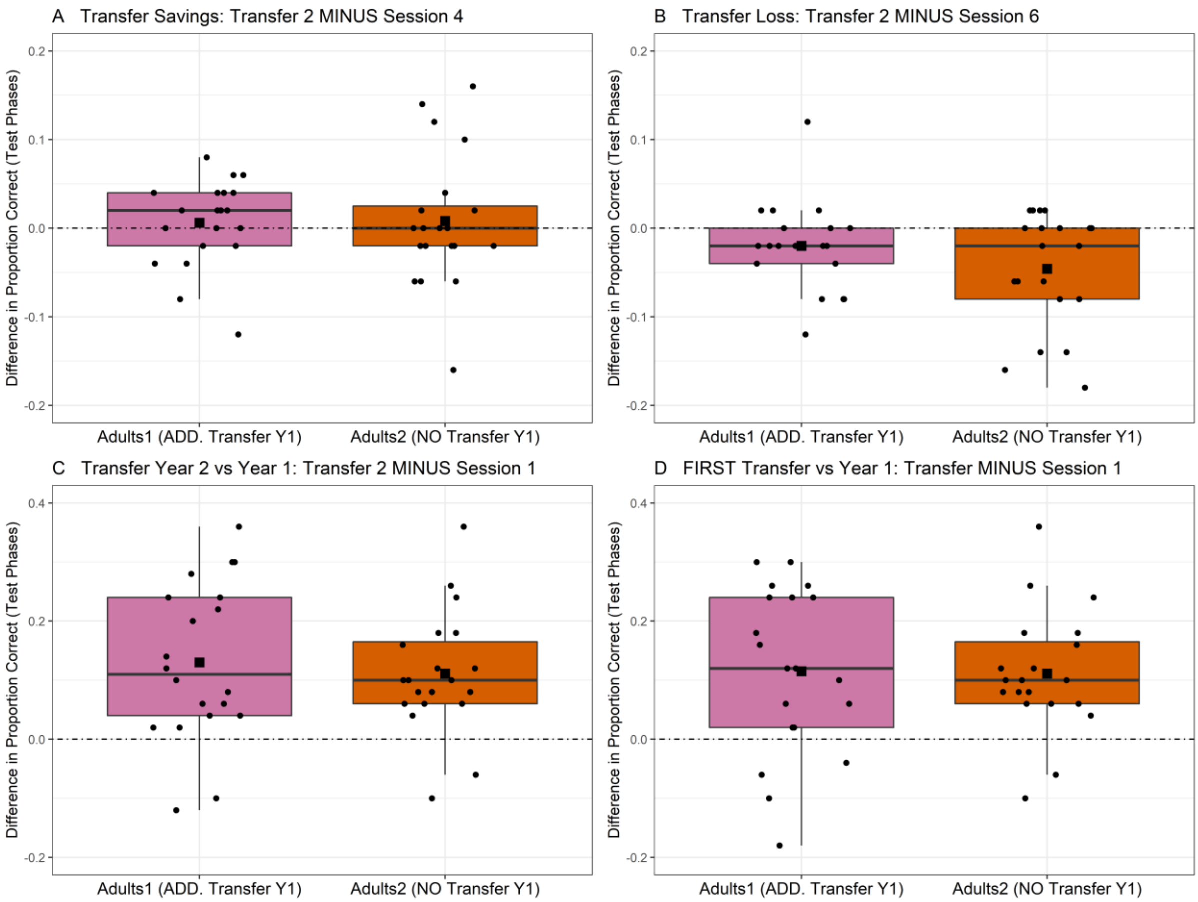
Transfer effect	Session levels	Interaction effect *Group*Session*
Transfer Savings	Session 4 vs. Transfer 2	$F(1, 38) = 0.01$, $p = .922$, $\eta^2_g < .01$, $BF_{10} = .32$
Transfer Loss	Session 6 vs. Transfer 2	$F(1, 38) = 2.01$, $p = .165$, $\eta^2_g < .01$, $BF_{10} =.63$
Transfer Year 2 vs. 1st Learning	Session 1 vs. Transfer 2	$F(1, 38) = 0.25$, $p = .618$, $\eta^2_g < .01$, $BF_{10} = .36$
First Transfer vs. 1st Learning	Session 1 vs. Transfer 2 (Adults 2), Session 1 vs. Transfer 1 (Adults 1)	$F(1, 38) = 0.01$, $p = .919$, $\eta^2_g < .01$, $BF_{10} = .33$

Note. All main effects of *Group*: $F(1, 38) \leq 1.73$, $p \geq .197$, $\eta^2_g \leq .04$, $BF_{10} \leq .81$.

Thus, both adult groups showed comparable transfer effects, independent of their session timing, i.e., independent of whether they had completed a transfer session in Year 1 already or participated in 7 vs. 8 AGL sessions before transfer was tested in Year 2. Consequently, Chapter III includes 7-year-olds alongside 5-year-olds and 6-year-olds for a more comprehensive perspective on repeated sequence learning in development, despite 7-year-olds completing an additional transfer session in Year 1 (equivalent to Adults 1).

Figure C.2

Transfer Effects for Adults 1 (pink) vs. Adults 2 (orange)



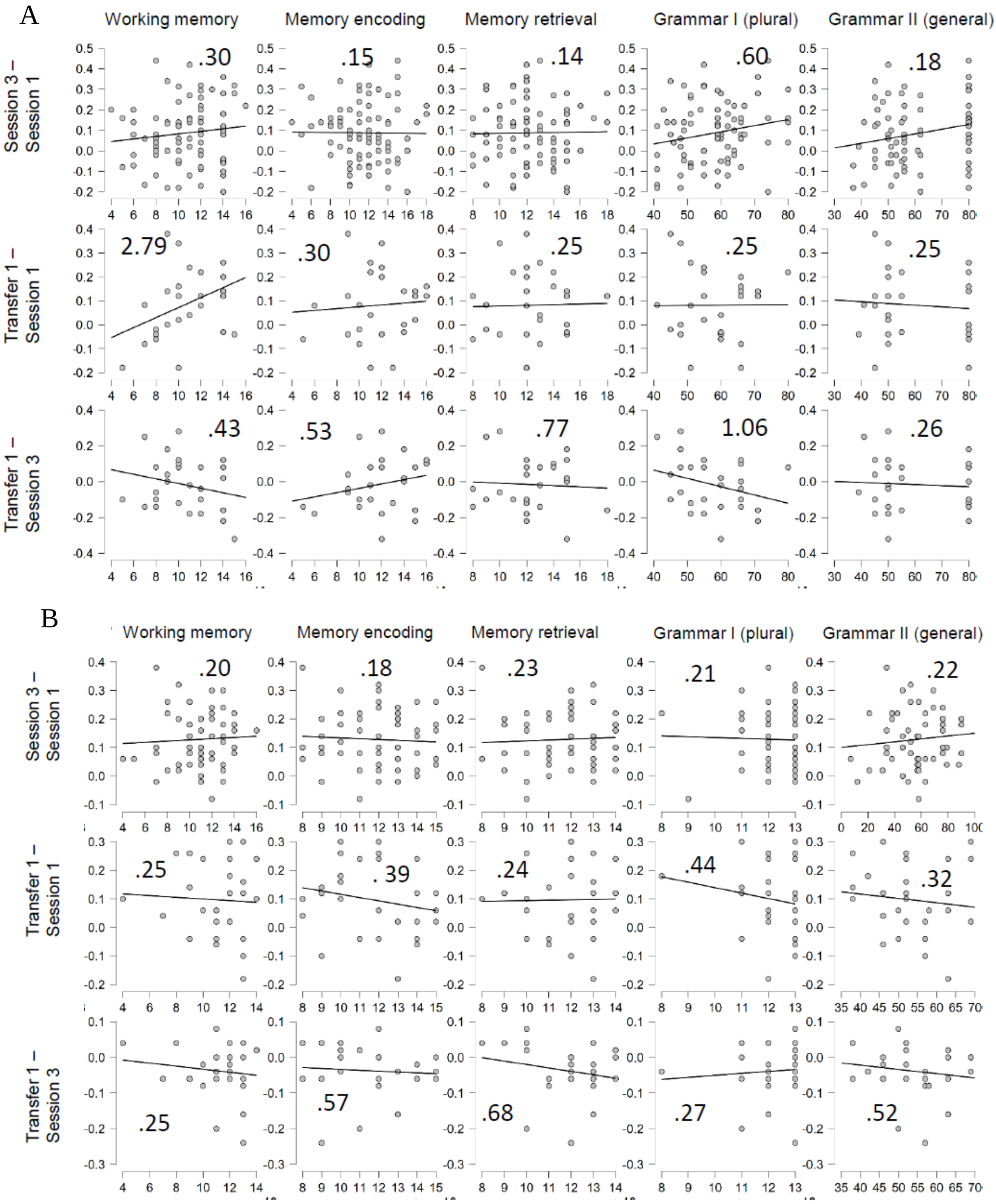
Note. Transfer Savings as difference in proportion correct responses of Transfer 2 and Session 4 (A), Transfer Loss as difference in proportion correct responses of Transfer 2 and Session 6 (B). Transfer as difference in proportion correct responses of Transfer 2 and Session 1 (C), and as difference in proportion correct responses of the first transfer session (Adults 1: Transfer 1, Adults 2: Transfer 2) and Session 1 (B) respectively. Boxplots for Adults 1 (pink) and Adults 2 (orange) with the groups' median indicated by a black line and the corresponding mean by a black square. Black dots represent single-subject data. The dotted lines mark no performance difference between the two compared sessions. ADD. = additional, Y1 = Year 1.

APPENDIX D:
Additional Analyses Chapter IV (Project 3)

Scatterplots: Correlations of AGL performance (Year 1) with cognitive skills (Year 1)

Figure D.1

Scatterplots for Cognitive Skills (x-axis) and AGL performance (y-axis) within Year 1



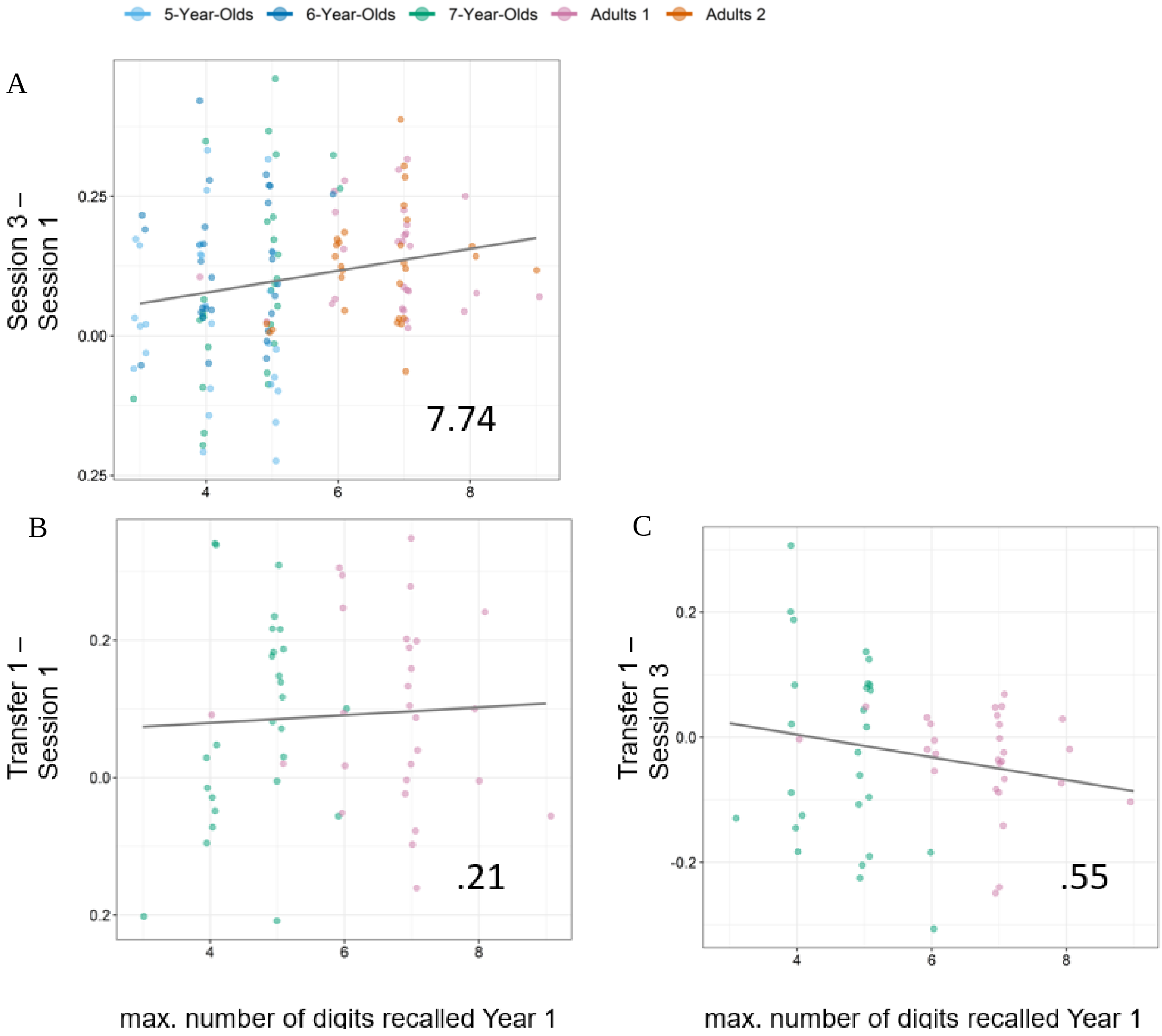
Note. Memory and language skills and AGL session differences plotted for children (A) and adults (B). Gray dots indicate individual subjects. Black lines denote linear regression. Numbers per plot indicate BF_{10}.

Scatterplots: Correlations of AGL performance (Year 1) with cognitive skills (Year 1)

(continued)

Figure D.2

Scatterplots for Working Memory (x-axis) and AGL performance (y-axis) within Year 1

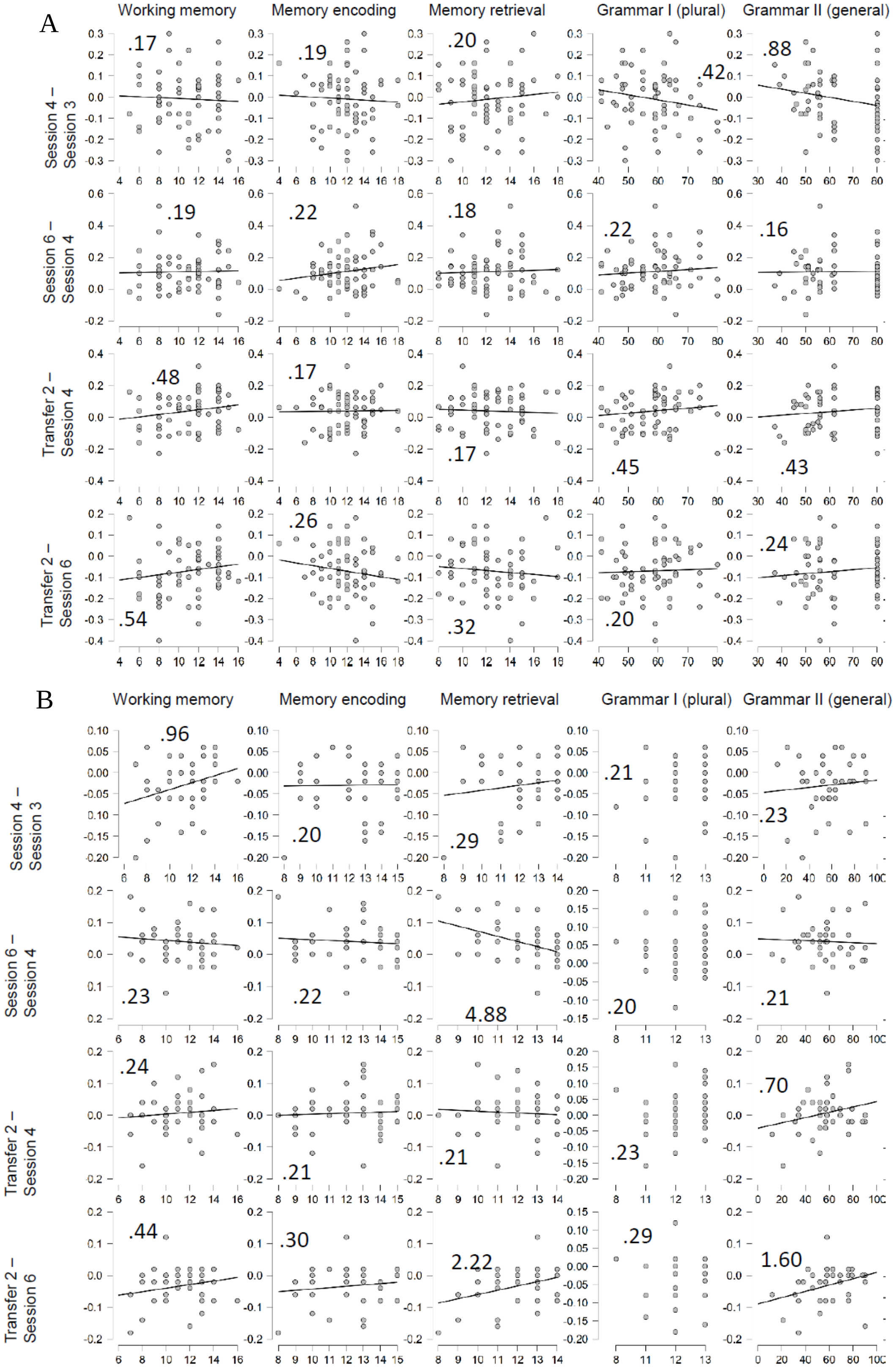


Note. Working memory as max. digit span recalled and AGL session differences (A: Learning Gains, B: Transfer Savings, C: Transfer Loss) plotted for the combined sample of children and adults. Colored dots indicate individual subjects. Gray lines denote linear regression. Numbers per plot indicate BF_{10}.

Scatterplots: Correlations of AGL performance (Year 2) with cognitive skills (Year 1)

Figure D.3

Scatterplots for Cognitive Skills Year 1 (x-axis) and AGL performance Year 2 (y-axis)



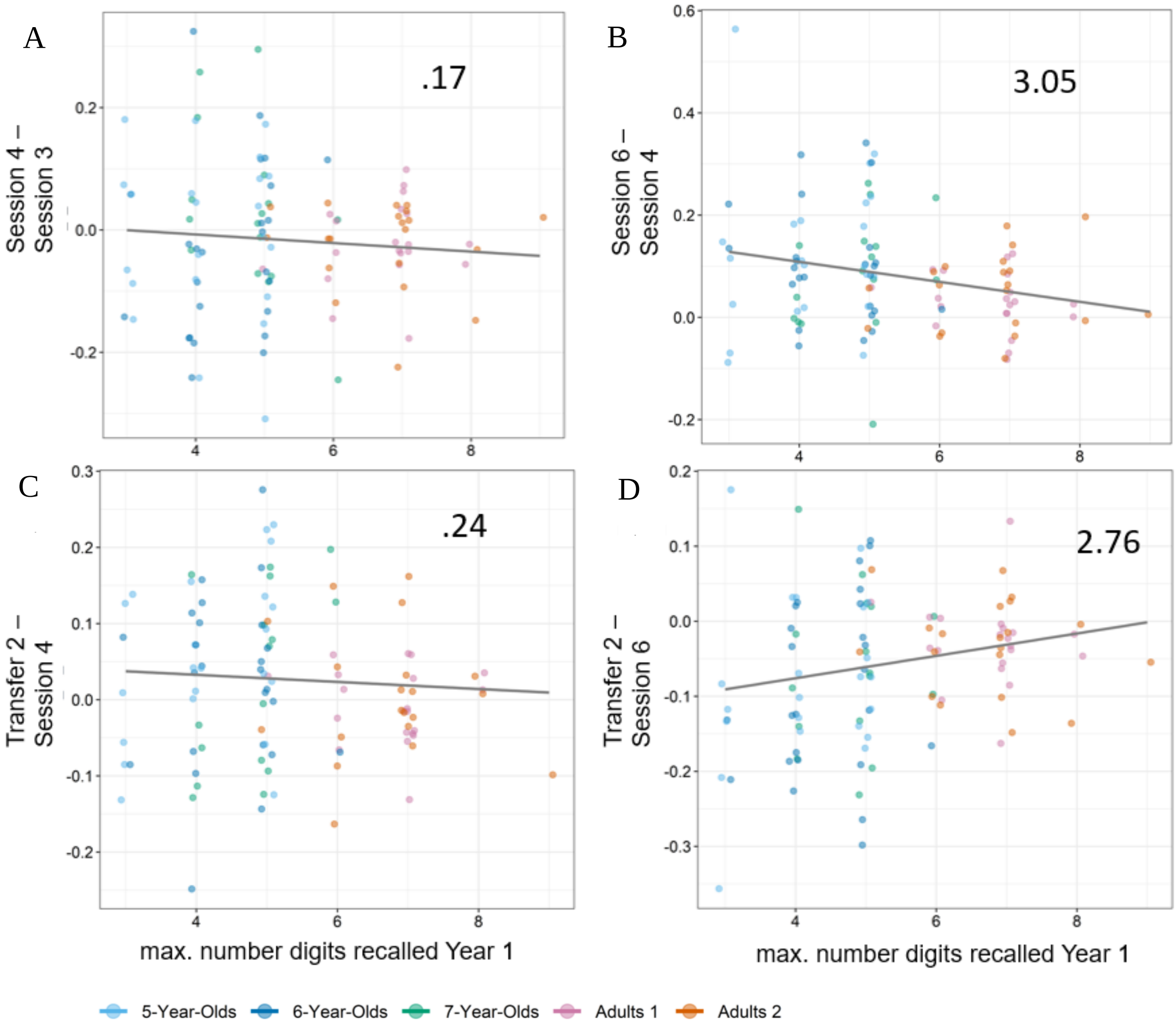
Note. Memory and language skills and AGL session differences plotted for children (A) and adults (B). Gray dots indicate individual subjects. Black lines denote linear regression. Numbers per plot indicate

Scatterplots: Correlations of AGL performance (Year 2) with cognitive skills (Year 1)

(continued)

Figure D.4

Scatterplots for Working Memory Year 1 (x-axis) and AGL performance (y-axis) Year 2

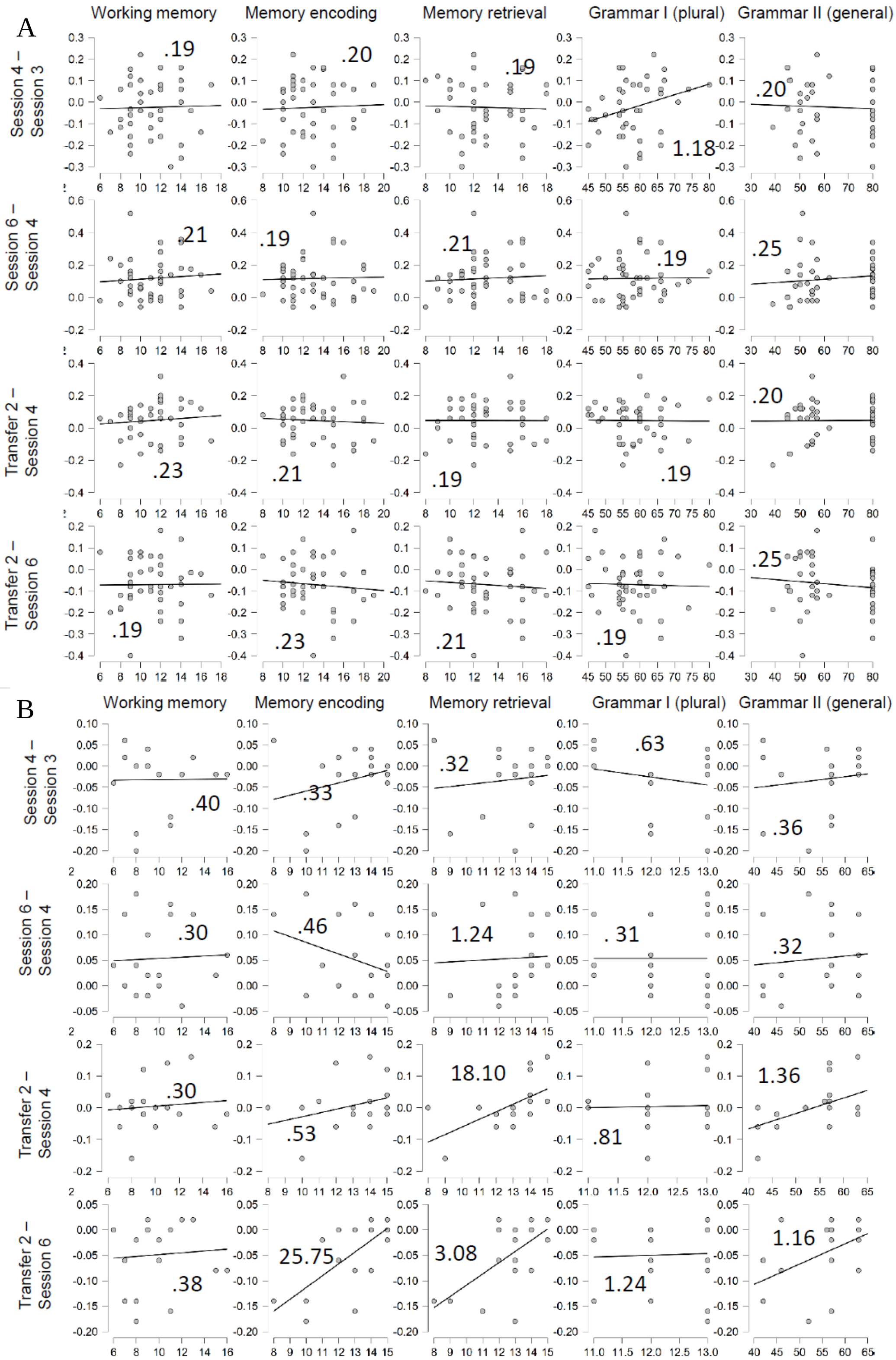


Note. Year 1 working memory as max. digit span recalled and Year 2 AGL session differences (A: Retention, B: Learning Gains, C: Transfer Savings, D: Transfer Loss) plotted for the combined sample of children and adults. Colored dots indicate individual subjects. Gray lines denote linear regression. Numbers per plot indicate BF_{10}.

Scatterplots: Correlations of AGL performance (Year 2) with cognitive skills (Year 2)

Figure D.5

Scatterplots for Cognitive Skills Year 2 (x-axis) and AGL performance Year 2 (y-axis)



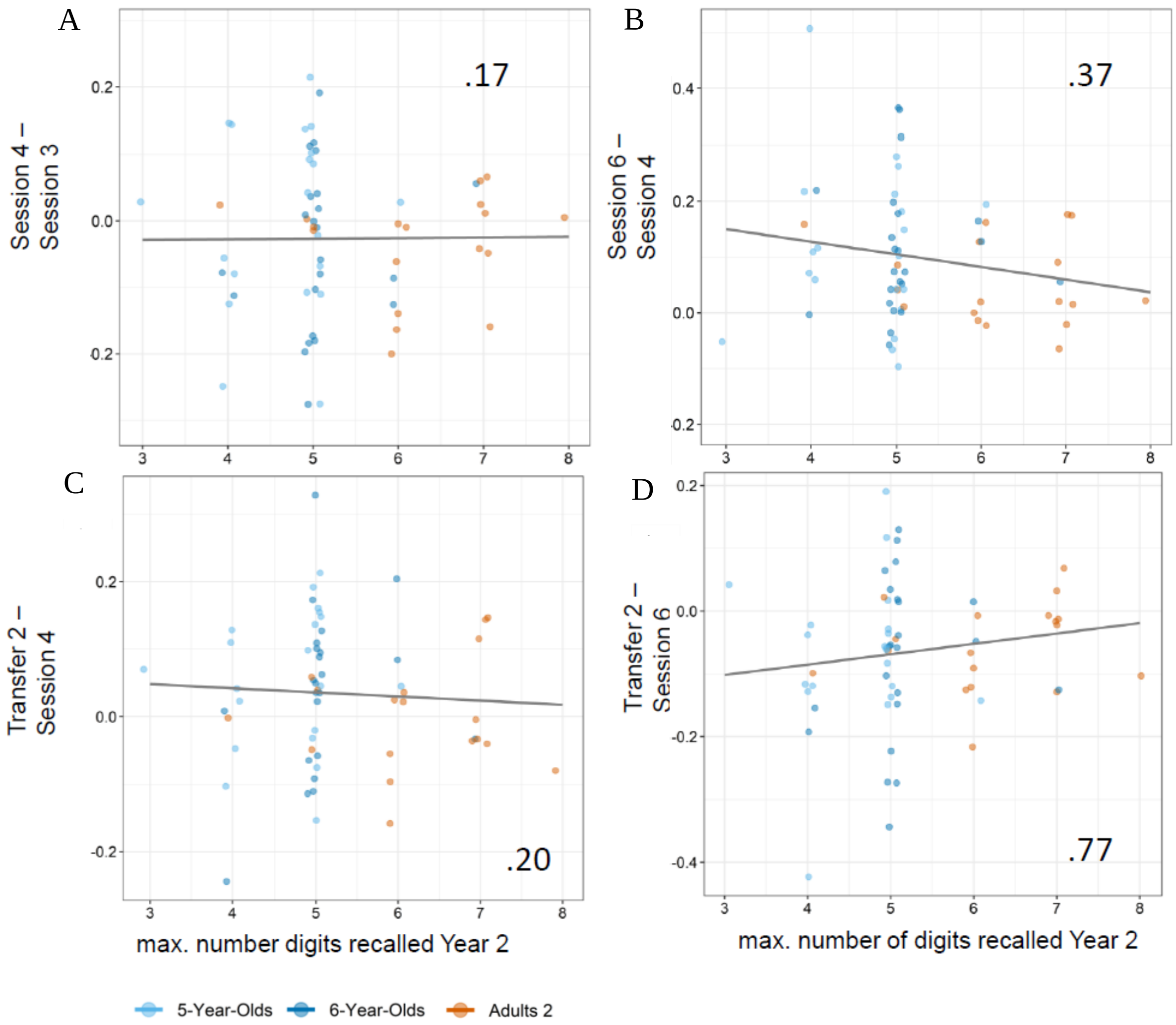
Note. Memory and language skills and AGL session differences plotted for children (A) and adults (B). Gray dots indicate individual subjects. Black lines denote linear regression. Numbers per plot indicate BF_{10}.

Scatterplots: Correlations of AGL performance (Year 2) with cognitive skills (Year 2)

(continued)

Figure D.6

Scatterplots for Working Memory in Year 2 (x-axis) and AGL performance (y-axis) in Year 2



Note. Year 2 working memory as max. digit span recalled and Year 2 AGL session differences (A: Retention, B: Learning Gains, C: Transfer Savings, D: Transfer Loss) plotted for the combined sample of children and adults. Colored dots indicate individual subjects. Gray lines denote linear regression. Numbers per plot indicate BF_{10}.

Psychometric Measures for cognitive skills:

Correlations across 1 year in children and adults with 2 available assessments

Table D.1

Correlations of Psychometric Measures for Cognitive Skills in Children and Adults

	Test-Retest Reliability (Session 1 ~ Session 4)	
	Children (5- & 6-year-olds, *n* = 43)	**Adults 1 (*n* = 18)**
German Grammar I (Plural)	.17 [-.24-.44]	.27 [-.22-.66]
German Grammar II (General)	-.21 [-.48-.10]	.89 [.72-.96]
Declarative Memory		
Encoding	.55 [.30-.73]	.39 [-.10-.72]
Retrieval	.34 [.05-.59]	.38 [-.11-.72]
Working Memory		
Raw Score	.55 [.29-.73]	.87 [.68-.95]
Max. Number of Digits recalled	.68 [.51-.79]	

Note. Correlations denote *Spearman's Rho* and were calculated on proportion correct of AGL test phases averaged across sessions and raw scores of psychometric assessments. [...] = 95% Confidence Interval (CI).

Table D.2

Separate Correlations in Adult Groups for Memory & Grammar Skills in Year 1 with AGL Parameters in Year 2

	Retention Year 1 to Year 2 AGL (Session 4 – Session 3)		**Learning Gains AGL** (Session 6 – Session 4)		**Transfer Savings AGL** (Transfer 2 – Session 4)		**Transfer Loss AGL** (Transfer 2 – Session 6)	
	Adults 1 (*n* = 20)	Adults 2 (*n* = 20)	Adults 1 (*n* = 20)	Adults 2 (*n* = 20)	Adults 1 (*n* = 20)	Adults 2 (*n* = 20)	Adults 1 (*n* = 20)	Adults 2 (*n* = 20)
German Grammar I (Plural)	.19	-.27	-.10	.07	-.18	***.45***	-.13	**.58***
German Grammar II (General)	.28	-.07	-.28	-.13	-.11	.20	.29	***.36***
Declarative Memory								
Encoding	-.19	.08	.08	-.17	-.08	.08	-.15	.27
Retrieval	<.01	.03	-.27	***-.32***	-.13	-.11	.20	.26
Working Memory Normalized Score	**.60 ***	.11	***-.38***	-.10	-.19	.08	.11	.24

Note. AGL = Artificial Grammar Learning Task, ***italic*** = $|r| \geq .30$, gray = $BF_{10} \geq 3$.

* corrected $p < .05$.

Acknowledgements

Diese Forschung wurde von der Deutschen Forschungsgemeinschaft mit Forschungsmitteln für Brigitte Röder gefördert (DFG Ro 2625/10-1).

Vielen Dank an alle, die mich in meinem Dissertationsprojekt und während der Promotionszeit unterstützt haben!

Ich möchte meiner Doktormutter Brigitte Röder, die mir während der fünf Jahre viel ermöglicht und zugetraut hat, für ihre wissenschaftliche Förderung danken. Vielen Dank auch an meinen Betreuer Patrick Bruns, der mir bei allen Fragen mit viel Geduld und der nötigen Portion Pragmatismus weitergeholfen hat. Danke auch an Ulf Liszkowski, Barbara Hänel-Faulhaber und Jutta Mueller als Mitglieder meines Promotionskomitees. Ich danke Ulrike Basten-Wissel, die mich bereits während des Bachelorstudiums für die Wissenschaft begeistert und mit ihrer Herangehensweise an psychologische Forschung inspiriert hat.

Vielen Dank an alle teilnehmenden Familien für ihre durchgängige Unterstützung trotz der Corona-Pandemie und die Bereitschaft, über viele Sitzungen und auch Zuhause an der Studie mitzuwirken. Ein großes „Danke“ an alle Studierende und technische Assistentinnen, ohne die meine Studien nicht durchführbar gewesen wären: Maike Fuchs, Alina Scheller, Kathrin Lambeck, Jelka Wöbke, Dagmar Tödter, Nicola Kaczmarek, Nele Westermann, Lena Gräfe, Lina Graumann, Luca Klapdohr, Leon Bauer und Phuc Ngyuen. Ich danke meinen Kolleg*innen Liesa Stange, Alexander Kramer und Veronika Zweckerl für hilfreichen Input im Projekt und praktische Hilfen bei technischen Umsetzungen, und allen weiteren Mitgliedern der Biologischen Psychologie und Neuropsychologie für den fachlichen Austausch. Danke an alle von ILMA, insbesondere an Andreas Weiß, Madita Linke, Alexander Kramer, Carolin Heitmann, Rashi Pant, Veronika Zweckerl, Valentina Marcazzan und Laura Kuhne, die auch in schwierigen Phasen immer da waren. Saving the best for last, den allergrößten Dank an meine beste Bürokollegin Liesa Stange, die mich durch alle Höhen und Tiefen begleitet hat und darüber zu einer lieben Freundin geworden ist!

Danke an meine Familie, meine Eltern und meine Schwester, die mir immer viel Verständnis entgegen gebracht und mich nach Kräften unterstützt haben. Am meisten Dank und Liebe geht an Philipp, der mir unglaublich viel Selbstvertrauen, Durchhaltvermögen und den nötigen Blick für alles Schöne im übrigen Leben geschenkt hat. Lieben Dank an alle weiteren wertvollen Menschen, die mich in der Zeit begleitet haben, v.a. an Mareike Klafka, die mir immer zugehört und neue Perspektiven eröffnet hat.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT
Institut für Bewegungswissenschaft
Institut für Psychologie

Erklärung gemäß ***(bitte Zutreffendes ankreuzen)***

- [ ] **§ 4 (1c) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**
- [x] **§ 5 (4d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich,

Daniela Schönberger (Vorname, Nachname),

dass ich mich an einer anderen Universität oder Fakultät noch keiner Doktorprüfung unterzogen oder mich um Zulassung zu einer Doktorprüfung bemüht habe.

Hamburg, 11.10.23
Ort, Datum

D. Sch
Unterschrift

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT
FÜR PSYCHOLOGIE UND
BEWEGUNGSWISSENSCHAFT

Institut für Bewegungswissenschaft
Institut für Psychologie

Eidesstattliche Erklärung nach ***(bitte Zutreffendes ankreuzen)***

☐ **§ 7 (4) der Promotionsordnung des Instituts für Bewegungswissenschaft der Universität Hamburg vom 18.08.2010**

☒ **§ 9 (1c und 1d) der Promotionsordnung des Instituts für Psychologie der Universität Hamburg vom 20.08.2003**

Hiermit erkläre ich an Eides statt,

1. dass die von mir vorgelegte Dissertation nicht Gegenstand eines anderen Prüfungsverfahrens gewesen oder in einem solchen Verfahren als ungenügend beurteilt worden ist.
2. dass ich die von mir vorgelegte Dissertation selbst verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und keine kommerzielle Promotionsberatung in Anspruch genommen habe. Die wörtlich oder inhaltlich übernommenen Stellen habe ich als solche kenntlich gemacht.

Hamburg, 11.10.23
Ort, Datum

D Sdz
Unterschrift