Open Research Online

The Open University's repository of research publications
and other research outputs

Hebb repetition learning in adolescents with intellectual disabilities.

Journal Item

oro.open.ac.uk

Hebb repetition learning in adolescents with intellectual disabilities

Lucy A. Henry [a],[*], Sebastian Poloczek [b], David J. Messer [a],[c], Rachel Dennan [a], Elisa Mattiauda [a], Henrik Danielsson [d],[e]

[a] Division of Language and Communication Science, City, University of London, UK
[b] Department of Psychology, Goethe University, Frankfurt, Germany
[c] Faculty of Wellbeing, Education & Language Studies, Open University, UK
[d] Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden
[e] Linnaeus Centre HEAD, Swedish Institute for Disability Research, Linköping University, Linköping, Sweden

ABSTRACT

Background: Hebb repetition learning is a form of long-term serial order learning that can occur when sequences of items in an immediate serial recall task are repeated. Repetition improves performance because of the gradual integration of serial order information from short-term memory into a more stable long-term memory trace.
Aims: The current study assessed whether adolescents with non-specific intellectual disabilities showed Hebb repetition effects, and if their magnitude was equivalent to those of children with typical development, matched for mental age.
Methods: Two immediate serial recall Hebb repetition learning tasks using verbal and visuospatial materials were presented to 47 adolescents with intellectual disabilities (11–15 years) and 47 individually mental age-matched children with typical development (4–10 years).
Results: Both groups showed Hebb repetition learning effects of similar magnitude, albeit with some reservations. Evidence for Hebb repetition learning was found for both verbal and visuo-spatial materials; for our measure of Hebb learning the effects were larger for verbal than vi-suospatial materials.
Conclusions: The findings suggested that adolescents with intellectual disabilities may show im-plicit long-term serial-order learning broadly commensurate with mental age level. The benefits of using repetition in educational contexts for adolescents with intellectual disabilities are considered.

1. Introduction

Hebb repetition learning (Hebb, 1961) is a form of incidental learning that occurs when sequences of items are repeatedly pre-sented in the same order. It involves a core learning mechanism that gradually transfers serial order information in short-term memory (e.g., phonological word forms) into stable long-term memory representations via repeated exposure, or repetition (Attout, Ordonez Magro, Szmalec, & Majerus, 2020). Procedural memory is believed by some to be the dominant system involved in Hebb repetition

learning (HRL) - at least in children - and is largely based on implicit processes (Attout et al., 2020). Procedural memory (Hsu & Bishop, 2014) is an important system for commonly repeated, well-known or 'automated' activities in different domains of everyday life, such as acquiring habits, learning implicit rules of grammar, learning new words, or skilled reading (e.g., Bogaerts, Szmalec, Hachmann, Page, & Duyck, 2015; Page & Norris, 2009).

Whilst there is an established literature on systems that support immediate 'online' storage and information processing, i.e., 'working memory' in individuals with non-specific intellectual disabilities (ID) (e.g., Henry, Messer, & Poloczek, 2018), there is no literature about HRL in these individuals. Consequently, we know little about this core mechanism that could be relevant to learning in several important domains for adolescents with ID such as grammar, vocabulary and reading (e.g., Attout et al., 2020; Bogaerts, Szmalec, De Maeyer, Page, & Duyck, 2016; Smalle, Page, Duyck, Edwards, & Szmalec, 2018), potentially acting as a powerful learning tool for this group.

The current study explored the nature and extent of HRL in adolescents with non-specific ID, comparing them with younger children with typical development (TD) matched for mental age. ID is a neurodevelopmental condition that begins in childhood, and is characterised by intellectual and adaptive functioning difficulties within conceptual, social, and practical domains. It is a common condition, with a prevalence of between 1% and 3% worldwide (McKenzie, Milton, Smith, & Ouellette-Kuntz, 2016; Patel, Cabral, Ho, & Merrick, 2020).

As HRL is a form of incidental long-term serial order learning that occurs when sequences of items in an immediate serial recall task are repeated (Hebb, 1961), it is distinct from short-term recall for this information (Mosse & Jarrold, 2010). HRL reflects the gradual integration of serial order information within short-term memory into a more stable and unified long-term memory trace (Bogaerts, Siegelman, Ben-Porat, & Frost, 2018). Thus, sequences maintained for immediate recall contribute to long-term memory traces, and long-term memory traces, in turn, help improve immediate serial recall when sequences are repeated (Oberauer, Jones, & Lewandowsky, 2015). HRL is usually measured using immediate serial recall tasks for letters, digits, syllables, words, spatial locations or nonsense pictures. The current study focused on immediate serial recall tasks for verbal (words) and visuospatial (nonsense pictures) materials in which some to-be-remembered sequences were repeated ('Hebb trials'), whereas other sequences were always novel ('filler trials'). Differences in serial recall between Hebb and filler sequences that emerge over trials provide an assessment of Hebb learning (Hebb, 1961).

The 'developmental approach' to intellectual disability suggests that cognitive abilities in those with non-specific ID should follow the same pattern of development as TD peers, albeit at a slower rate and perhaps reaching a lower asymptote (Burack et al., 2021). Consequently, one would expect similar HRL in groups with ID and TD matched for mental age. Limited support for the developmental approach comes from the finding that implicit learning tends not to differ between mental age-matched groups with ID and TD (Weiss, Weisz, & Bromfield, 1986), however, there is no directly-relevant evidence concerning HRL.

Mosse and Jarrold (2010) noted that preserved HRL in those with ID could have important educational implications, suggesting that knowledge about long-term implicit learning mechanisms for serial order could be applied in the classroom to improve educational outcomes. In their study of young people with Down syndrome, Mosse and Jarrold (2010) reported no group differences when they compared HRL in young people with Down syndrome and children with TD matched for verbal (non-verbal in Experiment 1) ability. This was despite well-documented relative difficulties with verbal short-term memory of individuals with Down syndrome (e. g., Jarrold, Purser, & Brock, 2006). These findings supported preserved HRL in young people with Down syndrome (i.e., mental age appropriate HRL, in accordance with the developmental approach), despite their lower verbal short-term memory performance.

To further extend understanding of HRL processes in adolescents with non-specific ID, we assessed HRL for both verbal and visuospatial materials, and compared them with a carefully mental-age matched comparison group of TD children. There are indications that young TD children show HRL in both verbal and non-verbal domains (e.g., Mosse & Jarrold, 2008; although see West, Vadillo, Shanks, & Hulme, 2018), as adults do (Couture & Tremblay, 2006; Page, Cumming, Norris, Hitch, & McNeil, 2006), supporting arguments for a common 'domain-general' learning mechanism (Couture & Tremblay, 2006; Mosse & Jarrold, 2008). Hitch, Flude, and Burgess (2009) suggested this mechanism could reflect the episodic buffer component of working memory, with item information stored, for example, in the phonological loop or visuospatial sketchpad (Baddeley, 2000). This debate focuses on adults (e.g., Johnson, Dygacz, & Miles, 2017; Sukegawa, Ueda, & Saito, 2019), so there remains uncertainty around domain-general mechanisms in children and adolescents. Furthermore, differences in HRL between verbal and visuospatial domains could emerge in individuals with ID, as they appear to have less well-preserved verbal short-term memory as opposed to visuospatial short-term memory (Henry et al., 2018; Lifshitz, Kilberg, & Vakil, 2016), and weaker verbal as opposed to visuospatial explicit long-term memory (Lifshitz-Vahav & Vakil, 2014). Consequently, comparisons between these domains can contribute to a better understanding of modality similarities or differences related to HRL.

Three research questions were addressed here. First, do adolescents with ID and children with TD matched for mental age both show HRL? We predicted that participants with ID would show HRL effects. This prediction was tentative, as no directly relevant previous evidence was available. We also predicted that children with TD (matched for mental age) would show HRL effects. This prediction was derived from the literature showing such effects are present in TD children from the age of four years (Archibald & Joanisse, 2013; Attout et al., 2020; Bogaerts et al., 2016; Smalle et al., 2016, 2018; West et al., 2018; Yanaoka, Nakayama, Jarrold, & Saito, 2019).

Second, are the HRL effects of similar magnitude in the two groups? We predicted that the Hebb effect would be comparable in both groups, consistent with a preserved learning mechanism in those with ID and supporting the developmental approach. This prediction was tentative given the lack of previous research.

Third, are the HRL effects of similar magnitude for both verbal and visuospatial materials? Given previous evidence in the literature for a domain-general process of HRL (Mosse & Jarrold, 2008), we predicted that this form of learning would be similar for visuospatial

materials as well as verbal materials (based on Mosse & Jarrold, 2010).

2. Method

2.1. Participants

The study involved 47 (26 females) adolescents with ID who had a mean mental age of 84.81 months (7:1 years) and a comparison group of 47 (25 females) children with TD, individually matched for mental age, who had a mean mental age of 85.04 months (7:1 years). There was no significant difference between the groups on mean mental age, $t(92) = -0.09$, $p = .93$ (see Table 1); nor did the variances differ using Levene's test, $F(1,92) = 0.008$, $p = .95$. Teachers confirmed participants had spoken English for at least two years at the time of testing to ensure they could understand the tasks and participate fully.

Adolescents with non-specific ID (where the biological cause for the ID has not been identified) were recruited. We incorporated the DSM-5 (American Psychiatric Association, 2013) approach to the definition of ID into participant selection to understand better the strengths and weaknesses in those with both cognitive and adaptive difficulties. DSM-5 emphasizes adaptive functioning as well as intellectual functioning (Patel et al., 2020) along with less emphasis on exact cut-off scores (Burack et al., 2021). The study was preregistered on the Open Science Framework (OSF) https://osf.io/gkpwh/ (dated 17.01.19), and a minor change to the inclusion criteria for the ID group was registered on 08.08.19 under Transparent Changes on the OSF (citation: osf.io/a5724). The change facilitated recruitment and thereby increased the sample size by relaxing the Vineland adaptive behaviour scores in the ID group from a standardised score cut-off of 79 to 85. The current study reports data relevant to the first set of pre-registered research questions; we also report reliability data relevant to the third pre-registered research question here and in the supplementary materials (data relevant to the second set of pre-registered research questions will be reported in a separate paper).

The 11- to 15-year-olds with ID were recruited from 27 mainstream secondary schools in England (Greater London, Hertfordshire, Yorkshire, Cambridgeshire and Nottinghamshire). Teachers identified eligible young people if they had ID and no other diagnoses such as autism or Down syndrome. Participants were excluded if, after testing, they did not have: 1) a score of between 40 and 79 on the Stanford-Binet abbreviated intelligence scales (SB-5: Roid, 2003); and 2) a standardized score of between 40 and 85 on the overall Adaptive Behaviour Composite (ABC) of the Vineland Adaptive Behaviour Scales (Vineland-3: Sparrow, Cicchetti, & Saulner, 2016), or a standardised score of between 40 and 85 on at least one of the core domains (Communication, Daily Living Skills, Socialisation). Note that we included participants with SB-5 IQ scores in the borderline (70−79), mild (55−69) and moderate (40−54) ID range *provided* they also showed evidence of adaptive difficulties. Of the 47 participants, 14 had borderline, 23 had mild and 10 had moderate ID in terms of their SB-5 scores.

In the group with TD, 4- to 10-year-olds were recruited from seven mainstream primary schools in England (Greater London and Yorkshire). Participating schools were comparable to the ID group on socio-economic status, as a higher proportion of students than the national average were from ethnic minority groups, spoke English as an Additional Language and were eligible for pupil premium funding. Teachers identified eligible children who did not have any special education needs, or diagnosed developmental conditions such as autism. Children were included if their mental age, based on a composite verbal and non-verbal score derived from the SB-5 matched that of one participant from the ID group; and if they had a standardised score on the SB-5 above 79.

All young people with ID who met inclusion criteria and had full data available from the first HRL session were included, provided there was an individual match on mental age within 4 months in the group with TD (note: only one mental age match differed by 4 months, and 32 were exact matches). Some participants (2 with ID; 6 with TD) were missing HRL data from the second session (and one further TD child was missing visuospatial data from session 2) because data collection was stopped due to Covid-19; there was no reason to suspect systematic bias so these participants were retained. Some participants were excluded (9 with ID; 11 with TD) who

Table 1

Mean scores, (SD) and [ranges of scores] on key study variables for adolescents with ID and children with TD matched for mental age.

Variables	Young people with ID (*n* = 47)	Children with TD (*n* = 47)	Group differences
Chronological age	13 yrs 4 m (15 m) [137–191 m]	7 yrs 2 m (17 m) [57–123 m]	$t(92) = 22.41$, $p < .001$
Mental age	7 yrs 1 m (13 m) [55–111 m]	7 yrs 1 m (13 m) [59–111 m]	$t(92) = -0.09$, $p = .93$
SB5 IQ	63.36 (10.39) [47–79]	98.66 (10.76) [85–121]	$t(92) = -16.18$, $p < .001$
Word List Recall: span	3.44 (0.53) [2.00–4.50]	3.30 (0.58) [2.25–4.25]	$t(92) = 1.19$, $p = .24$
Visual Sequential Memory: span	3.49 (1.06) [2–6]	3.38 (1.01) [2–6]	$t(92) = 0.50$, $p = .62$
Vineland: Communication	73.81 (8.20) [54–86]	–	–
Vineland: Daily Living Skills	83.13 (11.20) [48–105]	–	–
Vineland: Socialisation	83.30 (10.52) [60–109]	–	–
Vineland: ABC	78 (7.88) [59–94]	–	–

met the inclusion criteria, because they could not be matched.

Ethical approval was granted by the relevant university committee. Written informed consent from parents/guardians and written and verbal assent from participants were gained before testing.

2.2. Design

A mixed factorial quasi-experimental study was conducted with three within-subjects factors (Hebb repetition task list type (Hebb, filler), trial position (eight trials for each list type), and type of material (verbal, visuospatial)), and one between-subjects factor (group (ID versus TD)). No blinding was employed.

2.3. Materials

Standardized assessments of cognitive ability and adaptive functioning were administered. Measures of verbal and of visuospatial short-term memory were used to assign participants to an appropriate difficulty level of HRL (see Table 1).

2.3.1. Cognitive Assessment

The abbreviated version of the SB-5 (Roid, 2003) was used, consisting of two subtests: verbal knowledge and non-verbal reasoning skills. Manual-derived mental ages in months were used for matching. The SB-5 is suitable for individuals with ID and has high reliability (.91–.98) (Roid, 2003). Vocabulary, grammar and single-word reading were also assessed, but not reported here.

2.3.2. Adaptive functioning

Parents of participants in the ID group completed the Vineland-3 Domain-Level Parent/Caregiver Form (Sparrow, Cicchetti, & Saulnier, 2016) via a 20-minute telephone interview (N = 43); when this was not possible parents completed the questionnaire by themselves (N = 4). Standardised measures of adaptive functioning in three domains (communication, daily living skills, socialisation), as well as an overall adaptive behaviour composite (ABC), were derived. The Vineland-3 is a reliable assessment for individuals with ID (test re-test reliabilities from .73 to .92).

2.3.3. Short-term memory

Word List Recall from the Working Memory Test Battery for Children (WMTB-C: Pickering & Gathercole, 2001) assessed verbal short-term memory. Participants listened to word lists spoken by the experimenter before repeating the list in the same order. Word lists increased incrementally, beginning with lists of one word, with six trials for each list length; if four out of six lists were recalled correctly, the next list length was administered. Testing was discontinued if three out of six or fewer trials were correct in a block. To obtain 'sensitive' span scores, the longest list length for which 4 out of 6 trials were correct was identified, and extra credits of 0.25 for each list at the next higher list length correctly recalled were added to span scores (e.g., a full pass at list length 3 plus two correct trials at list length 4 gives a sensitive span score of 3.5).

Visual Sequential Memory from the Test of Memory and Learning (TOMAL-2: Reynolds & Voress, 2007) assessed visuospatial short-term memory. Participants were presented with left to right horizontally displayed sequences of nonsense visual stimuli for five seconds. These were then removed and immediately re-presented in a different order. Participants were instructed to "point to the drawings in the order you saw them on the page before". Sequence lengths increased incrementally with two trials per length. Testing was discontinued if the participant failed to recall any items in the correct order for two consecutive trials. Memory span scores were derived by taking the child's span score as the longest list length at which perfect recall in serial order was achieved.

2.3.4. HRL tasks

Given concerns about weak HRL effects in children (Archibald & Joanisse, 2013; Bogaerts et al., 2016; West et al., 2018; Yanaoka et al., 2019) and the reliability of HRL (Bogaerts et al., 2018; West et al., 2018), we aimed to maximize the sensitivity and reliability of our measures of HRL. No participants with chronological and mental ages younger than four years were included (see Yanaoka et al., 2019). The lengths of Hebb and filler tasks were chosen according to the participant's short-term memory ability to prevent floor and ceiling effects (Archibald & Joanisse, 2013; Hsu & Bishop, 2014; Siegelman, Bogaerts, & Frost, 2017; Smalle et al., 2016; West et al., 2018; Yanaoka et al., 2019). We used sensitive scoring of both item presence and serial position (e.g., Smalle et al., 2016). Non-overlapping item sets were used for Hebb and filler trials to minimise inter-list confusability (Johnson et al., 2017; Mosse & Jarrold, 2010; Smalle et al., 2016; Yanaoka et al., 2019). Finally, a similar serial order reconstruction recall method for all tasks was used that produces reliable HRL effects (Johnson et al., 2017).

All participants received two versions of HRL tasks on an iPad: a visuospatial task using non-nameable, unfamiliar 'nonsense drawing' stimuli; and a verbal task using easily nameable pictures of common objects with one-syllable names (similar to Archibald & Joanisse, 2013 and Hsu & Bishop, 2014). The verbal task involved simultaneously hearing each item's name and seeing its picture, with a presentation rate of one item every 1.5 s and a 0.5 s interval between items. The dual presentation method was used because, at recall, participants saw pictures of all items in the relevant item pool and were required to respond by touching the relevant pictures in serial order, so dual presentation facilitated cross-modal mapping. The visuospatial nonsense picture stimuli were shown without any sound.

There were 8 Hebb trials (the same to-be-remembered sequence was repeated 8 times), alternating with 8 filler trials (randomly generated novel sequences on each trial), making a total of 16 trials. The tasks started with a filler trial and alternated thereafter with

Hebb trials. Items were drawn from different item sets for Hebb and filler trials (i.e., stimuli for Hebb and filler trials were non-overlapping). There were two item-sets each of: 8 nonsense pictures (see Fig. 1a); and 10 one-syllable nouns illustrated as black and white line drawings (List 1 = dog, car, kite, chair, bell, ring, sun, fish, sock, house; List 2 = book, cat, bus, cup, bed, pear, comb, ball, duck, shirt) (see Fig. 1b). Choice of item sets was counterbalanced across participants and sessions.

In each HRL task, participants were shown, one at a time, a sequence of items. They then were shown an array of all items from the relevant item set, 10 items for verbal HRL task and 8 items for the visuospatial HRL task. The array was placed in the lower half of the screen, with the stimuli in two rows, and these were in a different random order on each trial. At the top of the response screen horizontal lines also appeared, the number of lines corresponding to the length of the list being recalled (this provided a cue about how many responses were needed, see Figs. 1a and 1b). To recall the target sequence, participants sequentially touched the images on the array screen. After each touch, a black circle appeared on the relevant line at the top of the screen to signify the selection. An item could be selected more than once, although target sequences never contained repetitions.

Both HRL tasks were introduced with practice trials using entirely different item sets. There were four trials in each practice block, two with short list lengths (two items for non-verbal tasks and three items for verbal tasks) and two with the same list length as the actual task (which were related to the assigned difficulty level).

Participants received a virtual coin for each trial completed in every task (the coin 'landed' inside a money bag with a 'ping') and a virtual gold trophy was given at the end of each task .

2.3.5. Hebb task allocation

There were short and long versions of each Hebb task (long versions had list lengths that were one item longer than short versions). Participants with Word List Recall memory spans of 3.5 or greater received the long verbal Hebb task (list length 6 items; $N_{ID} = 23$, $N_{TD} = 21$). Participants with Word List Recall memory spans of 3.25 or less received the short verbal Hebb task (list length 5 items, $N_{ID} = 24$, $N_{TD} = 26$). Participants with Visual Sequential Memory spans of 3 or greater received the long visuospatial Hebb task (list length 4 items; $N_{ID} = 38$, $N_{TD} = 38$). Participants with Visual Sequential Memory spans of 2 or less received the short visuospatial Hebb task (list length 3 items, $N_{ID} = 9$, $N_{TD} = 9$). This task allocation aimed to ensure there was room for performance improvement, with most children receiving supraspan list lengths (e.g., Archibald & Joanisse, 2013; Hsu & Bishop, 2014). One potential disadvantage of this method was that for participants with lower spans, serial reconstruction at recall could have been relatively more demanding because there were more non-presented items in the response array. However, the benefits of standardising the item set sizes and titrating difficulty levels for all participants were regarded as more important.

2.3.6. Nonsense drawing familiarisation

Before administering the visuospatial task, participants were pre-familiarised with the nonsense drawings through playing 'Snap'. A pack of 64 cards was dealt to the participant and experimenter, consisting of 16 nonsense pictures identical to those in the visuospatial task, each repeated four times. On each card, only one side showed a nonsense picture. The participant and experimenter took turns to turn over a card from their face-down pile and if the card matched the previous turned over card, the first player to say 'snap' won all turned over cards. When all the cards in the pack had been turned over, the player with the largest pile of cards won the game. This game was designed to give the participants a standardised experience to develop representations of the unfamiliar pictures.

2.3.7. Counterbalancing

Two HRL sessions were administered, up to two weeks apart, to maximise the amount of data collected (without tiring participants) and assess test-retest and split-half reliability. The sequences that participants received varied as a result of list length (short or long versions). Each session involved a verbal and visuospatial HRL task, counterbalanced across the two sessions. The items in each set of sequences were chosen via semi-randomized selection, but otherwise did not differ in format. Two parallel versions of each task were
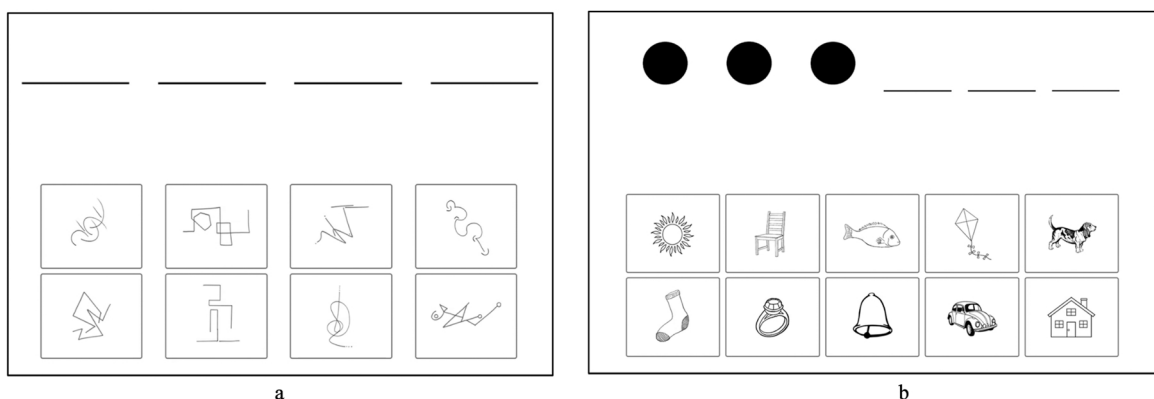

Fig. 1. a. Response array from the visuospatial Hebb repetition task. Here the participant has not started responding. Fig. 1b. Response array from the verbal Hebb repetition task. Here, the participant has made three responses as indicated by the black circles.

counterbalanced across participants: for session 2 the filler stimuli set and the Hebb stimuli set were reversed (i.e., the item set for Hebb sequences in session 1 then became the item set for filler sequences in session 2), and this was counterbalanced across participants. The specific version that the participant received first was randomised.

2.3.8. HRL Scoring

Given the known issues with obtaining reliable measures of inter-individual differences in Hebb learning (Bogaerts et al., 2018), credit for correct item and position information was given, to ensure that partial knowledge of the sequences was taken into account in the scores. As outlined by Kalm and Norris (2016), HRL involves not just learning item-position associations; there can also be partial knowledge of subsequences or chunks within the sequence. Item scoring captures this partial learning, even when exact serial order retention breaks down. Also, our response arrays for serial reconstruction contained more items than just those presented, so some scoring of item recall was necessary. For each item in a sequence, the participant was scored to take account of correctly recalling the item and its serial position as follows: 0 = 'no recall'; 1 = 'item recall (not in position)'; and 2 = 'item recall in position'. To further ensure that we did not underestimate the true effects of HRL, we calculated Levenshtein edit-distance metrics, "defined as the minimum number of edits needed to transform one string into another" (Kalm and Norris, 2016, p.112; edit distance was divided by list length and subtracted from 1 to derive a standardised metric). This gives credit for any similarities between the target sequence and the recalled sequence, making minimal assumptions about what is being learned. The Levenshtein scoring method and the relevant analyses were exploratory as they were not pre-registered (see the full analyses in supplementary materials, Section 4).

2.4. Procedure

Participants in both groups were assessed one-to-one at their schools during lesson time. Session lengths and the number of sessions were adapted to the participants' needs and school schedules. Most participants with ID completed the assessments in 90 min split across two sessions. During the first session the SB-5, short-term memory measures, verbal and visuospatial HRL tasks (and a reading measure if time allowed) were administered. The second session consisted of the second verbal and visuospatial HRL tasks, plus two language measures (and the reading measure as needed). The majority of children with TD completed the activities in three sessions of approximately 30 min each (session 1 included SB-5, short-term memory measures, reading task; session 2 included verbal and visuospatial HRL tasks and a language measure; session 3 included the second presentation of the verbal and visuospatial HRL tasks and another language measure). Certificates were provided to participants after the final session as a reward and primary school children also received stickers.

2.5. Analysis method

Recall performance on trials with repeated sequences (Hebb) was compared with recall performance on changing (filler) sequences. Performance was compared using all trials, for verbal and visuospatial materials in both sessions, in the ID and TD groups. Data were analysed with generalized linear mixed models (GLMM) to maximise sensitivity. Several studies have taken performance improvements from the first to the second half of Hebb trials relative to the performance change in filler lists as their measure of Hebb learning (e.g., Archibald & Joanisse, 2013; Mosse & Jarrold, 2008). The drawback of this 'halves' approach, as with any dichotomisation of continuous variables, is that information about changes within the first and second halves of trials is lost, resulting in a loss of power to detect effects (e.g., MacCallum, Zhang, Preacher, & Rucker, 2002). In other studies, separate regression analyses of performance on Hebb trials and filler trials were performed for each participant. The resulting regression slopes or gradients were entered into ANOVAs or ANCOVAs (e.g., Bogaerts et al., 2015; Hsu & Bishop, 2014).

However, an alternative approach to deal with the multilevel structure of trials being nested within participants is to analyse the data with generalized linear mixed models (GLMM) instead of performing separate consecutive analyses. Similar to the regression approach, changes in recall across trial position are modelled without any information loss about trial position. A key benefit of GLMM is that data from Hebb trials, filler trials, and from all participants are analysed in a joint model. A positive interaction effect between list type (Hebb vs. filler trials) and trial position (1 through 8) represents the degree of Hebb learning, as it captures recall improvement over Hebb trials in comparison to filler trials. The mixed GLMM models can include both this fixed effect of list type x trial position, representing the average effect of HRL, and the random effect of list type x trial position, allowing for inter-individual differences in the degree of HRL. In mixed effect models the residuals of the random effects reflect how an individual differs from the group mean (fixed effects). Including a random intercept additionally captures individual differences in recall performance on the first Hebb and filler trials. Bogaerts and colleagues (Bogaerts et al., 2016, 2018) introduced mixed logit models to analysing the development of Hebb learning with recall of items in the correct position as dependent, binary variable (for a similar approach see Yanaoka et al., 2019).

The information from scoring HRL on each of the three to six items constituting a trial was ordinal with three levels (0−2). We modelled the scoring of HRL using a cumulative logit model with proportional odds assumed. This means that only one effect per predictor (material, list type, trial position + list x position interaction) is estimated and the effect on the transition from no recall to recall is assumed to be the same or proportional in terms of odds as on the transition from recall not in position to recall in position. The resulting models were complex because various fixed effects for group-level experimental effects and random effects for individual differences had to be estimated. Therefore, only the interaction effect of list and position was included to capture Hebb learning. Consequently, further two-way interactions or even three-way interactions to test for differences in the degree of Hebb learning between participant groups or materials were not included in a single model. Increasing the model complexity further would increase the risk of problems in model estimation. Instead, four separate models were set up to address the specific research questions (compare

preregistration: https://osf.io/gkpwh/): models for participants with ID and TD, respectively, that pooled data across material and assumed the same degree of Hebb learning for verbal and visuospatial material; and models for verbal vs. visuospatial material, respectively, to test differences due to material that included data from both groups, but allowed for individual differences in Hebb learning by including the relevant random effects. All GLMMs for ordinal dependant variables were performed with MLwiN 3.02 (Charlton, Rasbash, Browne, Healy, & Cameron, 2018) using MCMC estimation, with 100000 iterations and thinning to 5000 estimates from R with the R2MLwiN package (Zhang, Parker, Charlton, Leckie, & Browne, 2016). On publication, full data and R scripts will be made available on the Open Science Framework.

2.6. Reliability of the HRL tasks

The test-retest correlations of the Hebb-learning residuals were: $r_{tt} = 0.47$ ($p < .001$) for the model averaging across material (RQ1, RQ2); $r_{tt} = 0.63$ ($p < .001$) for the verbal material; and $r_{tt} = 0.69$ ($p < .001$) for the visuospatial material (both relating to RQ3), resulting according to the Spearman-Brown formula in split-half-reliabilities of .64, .77, and .82, respectively.[1] The four correlations between HRL assessed with verbal and visuospatial materials (in each group) ranged from $r = 0.45$, 95% CI [.26, .60] to $r = 0.59$, 95% CI [.44, .71]. (See supplementary materials, Section 1, for further details).

3. Results

Concerning the first research question (do adolescents with ID and children with TD matched for mental age show HRL effects?), the plots in Fig. 2 suggest that HRL was shown in both groups. On the y-axis, the proportion correct from the combined scores for items being recalled, but not in position (0.5), and items being recalled in correct position (1.0), are displayed (for further descriptive results and a discussion of whether there were ceiling and floor effects present, see supplementary materials, Section 2). The figure suggests that HRL occurred in both groups, as a higher proportion of correct answers were given for Hebb sequences. There also appeared to be a separation between Hebb and filler sequences as the trials progressed in the verbal task. In line with previous research, involving TD children, Hebb trials showed maintenance of performance whilst filler trials tended to show declines (e.g., Archibald & Joanisse, 2013; Mosse & Jarrold, 2008).

To test these observations formally, GLMMs were run. Two separate models were used, one for adolescents with ID and one for children with TD. Both sets of results include random intercepts parameters for participants to capture inter-individual differences in how well participants remembered sequences, and random slopes for material and trial position. This allowed investigation of individual differences between verbal and visuospatial materials and for differences in how memory performance develops across trials. Crucially for the first research question, the random slope for the interaction of list type (Hebb vs. filler) and trial position (1 through 8) was included to allow for individual differences in how much participants benefited from the repeated presentation of Hebb lists. The fixed effects, and therefore the group level effects, of both models are displayed in Table 2. In ordinal models, instead of one intercept, multiple intercepts are estimated for each transition between categories. The two intercepts were comparable across groups indicating that the first trials of the experimental sessions (position value = 0; i.e. the aggregate for the 8 first trials of Hebb/ filler, verbal/visuospatial, and 1st/2nd sessions) were of comparable difficulty for both groups. This implies that both groups could start Hebb learning with, on average, comparable performance. However, the intercepts do not inform us about whether within-group differences, for example due to material, existed or not. In fact, the general effect of material was significant in both samples (ID: $\beta = 0.41$, 95%CI [0.33, 0.48], $p < .001$; TD: $\beta = 0.42$, 95%CI [0.32, 0.51], $p < .001$): participants were more likely to recall verbal sequences than visuospatial sequences, even though task difficulty was adjusted to the differing memory spans for these materials.

Of key interest for the first research question was the interaction between list type and trial position to identify an effect of Hebb learning. This interaction effect was significant in both groups (ID: $\beta = 0.05$, $p < .001$; TD: $\beta = 0.09$, $p < .001$) indicating that participants in both groups significantly benefitted from the repeated presentation of Hebb sequences compared to filler sequences. The non-significant main effect of trial position (ID: $\beta = 0.00$, $p = .71$; TD: $\beta = -0.01$, $p = .29$) suggests that recall performance did not improve with trial position overall. To examine whether the list x position interaction was driven by performance improvements on Hebb trials or by decreasing performance on filler trials, exploratory analyses were run on models testing the position effect either only on Hebb or only filler trials. Recall significantly improved on Hebb trials for both groups and materials (ID – verbal: $\beta = 0.04$, 95%CI [0.01, 0.08], $p = .02$; TD – verbal: $\beta = 0.09$, 95%CI [0.03, 0.15], $p = .002$; ID – visuospatial: $\beta = 0.05$, 95%CI [0.01, 0.09], $p = .02$; TD – visuospatial: $\beta = 0.06$, 95%CI [0.03, 0.10], $p < .001$) and significantly deteriorated on filler trials with verbal material (ID: $\beta = -0.10$, 95%CI [−0.13, −0.07], $p < .001$; TD: $\beta = -0.16$, 95%CI [−0.19, −0.13], $p < .001$) and did not change with visuospatial material, (ID: $\beta = -0.00$, 95%CI [−0.03, 0.03], $p < .92$; TD: $\beta = -0.03$, 95%CI [−0.06, 0.00], $p = .06$).

To answer the second research question (are HRL effects comparable for both groups?), the 90% credible intervals (CI) for the interaction between list type and trial position in each group were computed to determine whether they overlapped. As the interaction effect for Hebb learning was 0.05 with a 90% CI of 0.033–0.072 in the ID group and 0.09 with a 90% CI of 0.071–0.109 in the TD group, the CIs were overlapping, however more precisely, the CIs were touching and only barely overlapping. Therefore, although there was

---

[1] Reliability represents how consistently inter-individual differences can be measured. In mixed effect models the residuals of the random effects reflect how an individual differs from the group mean (fixed effects). To derive reliability estimates for Hebb learning, we therefore estimated separate GLMMs for session 1 and session 2, extracted the residuals of the random effects for the list type x position interaction (indicating Hebb learning) from each model and computed retest-reliabilities (for further details see supplementary materials, Section 1).
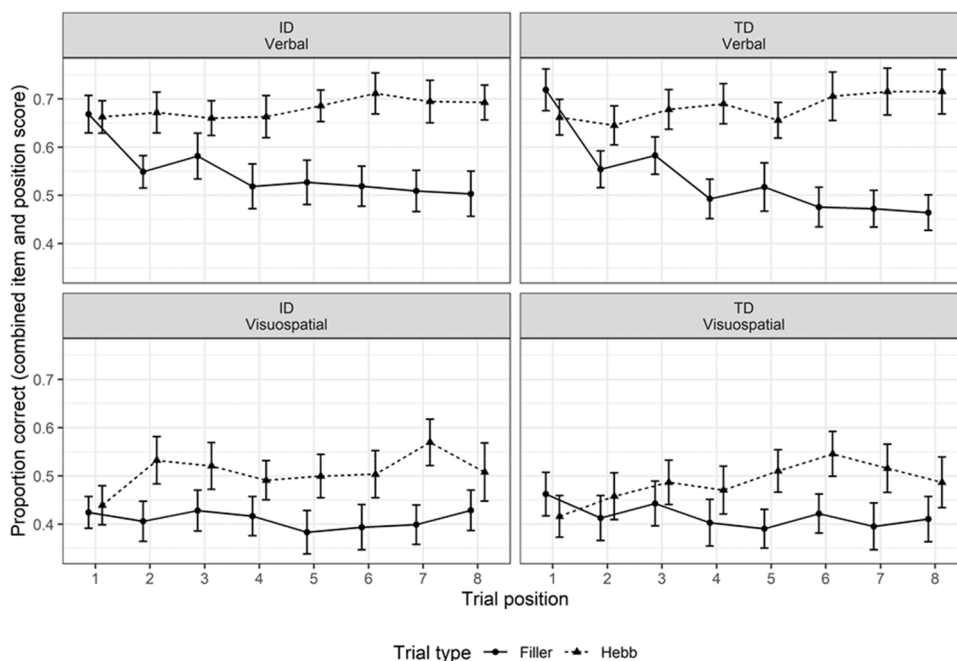
Fig. 2. Recall performance across trials for filler lists and repeated Hebb lists. Data are split by experimental groups (ID vs. TD) and material (verbal vs. visuospatial) and pooled across both experimental sessions (some data are missing for session two; n = 2 ID and n = 6–7 TD).

Table 2
GLMMs on recall performance across trial positions for Hebb vs. filler lists to address research question 1 for the ID and TD groups.

Fixed Part	ID			TD		
	β	95% CI	p	β	95% CI	p
Intercept 1 (no vs. item recall)	1.08	[0.89, 1.26]	< 0.001	1.15	[0.96, 1.33]	< 0.001
Intercept 2 (item vs. position recall)	-0.67	[− 0.93, − 0.42]	< 0.001	-0.67	[− 0.92, − 0.42]	< 0.001
Material (Verbal = +1 vs. Visuospatial = −1)	0.41	[0.33, 0.48]	< 0.001	0.42	[0.32, 0.51]	< 0.001
List (Hebb = +1 vs. Filler = −1)	0.16	[0.10, 0.22]	< 0.001	0.03	[− 0.03, 0.09]	.30
Trial Position (1st = 0, 2nd = 1 …)	-0.00	[− 0.02, 0.02]	.71	-0.01	[− 0.04, 0.01]	.29
List x Position (Hebb effect)	0.05	[0.03, 0.08]	< 0.001	0.09	[0.07, 0.11]	< 0.001
Random Part						

Random effects for random participant intercepts, for material, position, and for the list x position interaction (individual differences in Hebb learning), plus all covariances between these effects were included in the model and are displayed in the supplementary materials, Section 3 (Table S4).

Note. Sample sizes for the ID group: $N_{Participants} = 47$, $N_{Trials} = 2944$, $N_{Items} = 13680$; for the TD group: $N_{Participants} = 47$, $N_{Trials} = 2800$, $N_{Items} = 13024$

no robust evidence that HRL differed between groups, the degree of overlap was only just discernable, so some uncertainty remains regarding this result.

To answer the third research question (do children in both groups show a similar magnitude of HRL for both verbal and visuospatial materials?), two further GLMMs were conducted. The data were split by material type, but pooled across participant groups. Apart from the random intercepts for participants, the random slope for the list type x trial position interaction was included to allow for individual differences in Hebb learning. Results are displayed in Table 3. The crucial fixed effects are the interaction of list type (Hebb vs. filler) and trial position. The interaction effects were significant and positive for verbal stimuli (β = 0.100, 90%CI [0.081, 0.118], $p < .001$) as well as visuospatial stimuli (β = 0.034, 90%CI [0.019, 0.050], $p < .001$) indicating that for both types of materials Hebb learning was found, hence, children's recall benefitted from the repeated presentation of Hebb lists. As the 90% CIs were not overlapping, the HRL effect was more pronounced for verbal than visuospatial materials.

We also ran exploratory analyses using standardised Levenshtein distance scores (which were highly correlated with our combined item + position scores, $r = 0.94$, $p < .001$). Of particular interest were fixed effects for the list type x trial position interaction representing the average Hebb learning effect. Irrespective of the scoring method, the parameter estimates were the same for the first two decimal numbers in the analyses run by group: ID: β = 0.05, $p < .001$; TD: β = 0.09, $p < .001$. The results were also similar for the analyses run by material: verbal β = 0.09/0.10, $p < .001/< .001$; visuospatial: β = 0.03/0.03, $p = .003/< 0.001$ with the models based on standardised Levenshtein distance or on combined item + position scores, respectively. (See supplementary materials,

Table 3

GLMMs conducted on recall performance across trial positions for Hebb vs. filler lists for verbal and visuospatial stimuli to address research question 3.

Fixed Part	Verbal			Visuospatial		
	β	95% CI	p	β	95% CI	p
Intercept 1 (no vs. item recall)	1.74	[1.58, 1.9]	< 0.001	0.49	[0.33, 0.65]	< 0.001
Intercept 2 (item vs. position recall)	-0.27	[− 0.47, − 0.08]	.007	-1.04	[− 1.24, − 0.85]	< 0.001
List (Hebb = +1 vs. Filler = −1)	0.10	[0.05, 0.16]	< 0.001	0.09	[0.02, 0.16]	.007
Trial Position (1st = 0, 2nd = 1 …)	-0.03	[− 0.05, − 0.01]	.01	0.02	[0.01, 0.04]	.008
List x Position (Hebb effect)	0.10	[0.08, 0.12]	< 0.001	0.03	[0.02, 0.05]	< 0.001
Random Part						

Random effects for random participant intercepts, for position (only the model for verbal material), and for the list x position interaction (individual differences in Hebb learning), plus all covariances between these effects were included in the model and are displayed in the supplementary materials, Section 3 (Table S5).

Note. Sample sizes for verbal condition: $N_{Participants}$ = 94, N_{Trials} = 2880, N_{Items} = 15792; for visuospatial condition: $N_{Participants}$ = 94, N_{Trials} = 2864, N_{Items} = 10912

Section 4, for further details.).

4. Discussion

This first study of HRL in adolescents with non-specific ID showed, as tentatively predicted, a significant HRL effect in this group. Children with TD (individually matched for mental age) also showed significant HRL, supporting previous literature (Archibald & Joanisse, 2013; Attout et al., 2020; Bogaerts et al., 2016; Smalle et al., 2016, 2018; West et al., 2018; Yanaoka et al., 2019). The magnitude of the Hebb effect was similar in the two groups, although there was a degree of uncertainty here as the credible intervals for the groups were only just touching. As predicted, there were significant HRL effects for both visuospatial and verbal materials, although contrary to predictions from previous research (e.g., Mosse & Jarrold, 2010), the magnitude of HRL across these domains for our measure of Hebb learning was not equal, with larger effects for the verbal task. Although not part of our pre-registration, additional exploratory analyses using standardised Levenshtein distances as an alternative recall measure confirmed these findings and revealed a strong relationship between this measure and our original scoring method.

The findings broadly support the developmental approach (Burack et al., 2021) that adolescents with non-specific ID show HRL commensurate with their mental age level, implying that implicit long-term serial-order learning processes could be a relative strength in this group. Mosse and Jarrold (2010) similarly found that performance on HRL in their participants with Down syndrome was comparable to a group with TD matched for mental age, even though their group with Down syndrome showed weaker verbal short-term memory. However, the current findings were somewhat less supportive of the developmental approach than those of Moss and Jarrold (2010). This is because the credible intervals for the size of the Hebb repetition effect between the groups were touching rather than fully overlapping, suggesting that there could be a tendency for the magnitude of HRL to be lower in the ID group. In addition, verbal short-term memory in the present study did not differ between the groups with ID and TD (despite such findings often being reported: Henry et al., 2018; Lifshitz et al., 2016). This could be because the individual mental-age matching was very close, or because our inclusion criteria for ID, which included both cognitive and adaptive measures, differed from previous research.

Further exploratory analyses on the Hebb sequences showed increases in performance over trials for both types of material (in both groups). The filler sequences showed declines in performance (for both groups) in the verbal task, although there were no performance changes over trials for the visuospatial task. For the verbal filler sequences, the deterioration in performance might reflect a build-up of proactive interference or fatigue (Archibald & Joanisse, 2013; Bogaerts et al., 2016; Mosse & Jarrold, 2008). Interference is perhaps more likely; as trials progress, more confusion could occur on filler sequences because there are increasingly more previous sequences and individual items from the filler item pool that are 'partially' activated. By contrast, for items in the Hebb sequence, serial order and item information becomes repeatedly strengthened over trials, with less interference from items that are not in the Hebb sequence, leading to long-term learning (Hitch et al., 2009). In the visuospatial task, there was no clear evidence for such interference on filler trials. This could be because interference was not present or did not show in the data due to floor effects. Alternatively, it could be because the visuospatial items were not well-established in long-term memory, so were less likely to be partially activated and cause confusion as filler trials progressed. This latter point is interesting, as HRL could feasibly be conceptualized as just the Hebb learning element, rather than the interaction between learning on Hebb trials versus interference on filler trials. The suggestion of interference for verbal and not visuospatial filler trials could indicate that the precise conceptualization and operationalization of HRL for different types of materials may be more important than suspected. It is also important to note that because the current study employed separate Hebb and filler pools as item sets, interference effects could vary in scale and magnitude if all items were drawn from the same set, an issue that could be explored in future research.

The current findings supported the domain generality of HRL reported in both children and young people with Down syndrome (e. g., Bogaerts et al., 2016; Mosse & Jarrold, 2008, 2010), as HRL effects were significant for both visuospatial and verbal materials. Further, the HRL residuals taken from the verbal task correlated with the HRL residuals taken from the visuospatial task: this shared variance points to at least some domain generality. However, the magnitude of the HRL effect as indexed by the list type x trial position interaction was greater for verbal than visuospatial materials, whereas other studies have reported equal effects (Mosse & Jarrold,

2010), larger visuospatial effects (Bogaerts et al., 2016) or no visuospatial effects (West et al., 2018). It is not clear why findings differ, but task differences could be a possibility (e.g., static versus dynamic visuospatial tasks), as well as the problems of exactly matching the difficulty of tasks in different domains. There was some evidence that children with a higher starting performance level showed a tendency for higher HRL effects, so we cannot rule out the possibility that equating task difficulty would remove the differences in HRL between verbal and visuospatial materials. Thus, employing even more stringent difficulty level titration and including both static and dynamic visuospatial tasks would help to evaluate, in future research, whether HRL is equal for verbal and visuospatial materials. Larger scale studies would also enable the statistical models to take into account both group and material type in the same analysis, as more data points are required to ensure such complex models are stable. Overall, therefore, there was evidence to support a domain general mechanism for HRL, but further research is needed to explore this area, perhaps attempting to identify whether or not there are different serial order learning mechanisms across domains that share common features (e.g., Logie, Saito, Morita, Varma, & Norris, 2016) or whether commonalities between verbal, visual and spatial short-term memory imply domain-general serial order learning mechanisms (Hurlstone, Hitch, & Baddeley, 2014).

It has been argued that HRL "draws on the same memory processes responsible for representing and learning serial-order information in the service of language acquisition" (Bogaerts et al., 2015, p.107). Thus, serial-order HRL could boost the long-term acquisition of phonological sequences, which, in turn, are important for acquiring new vocabulary (Archibald & Joanisse, 2013; Mosse & Jarrold, 2008; Page & Norris, 2009; Smalle et al., 2018) and reading (Attout et al., 2020; Bogaerts et al., 2016); although relationships with language measures are not always found (Hsu & Bishop, 2014). Mosse and Jarrold (2010) discussed the educational implications of HRL, suggesting that indirect associations and learning opportunities afforded by Hebb processes may be more successful than direct and explicit instructional approaches using single teaching sessions. Further research could explore educational applications for HRL within real world learning settings, perhaps involving games and strategies that employ implicit, multiple presentations approaches or individualised programmes using HRL principles to improve the acquisition of specific vocabulary items for individual learners.

However, one key difference between HRL paradigms and real world language acquisition is that real world learning may have greater spacing and interspersed distractions between repetitions. Teachers often use repetition of vocabulary items to enhance word learning and this repetition could be distributed across one or more lessons. Some evidence suggests HRL is resilient to distraction at both encoding and retrieval. Oberauer et al. (2015) required both processing *and* storage within Hebb tasks, effectively making the immediate serial recall task a 'complex' memory span task. Participants were presented with Hebb and filler sequences in the usual way, but had to make judgements (about the sizes of pictured objects) after the presentation of each item in the sequence (or make these judgements at recall - interspersed between the recall of each to-be-remembered item in the sequence). Surprisingly, this degree of distraction did not minimise HRL effects in adults (Oberauer et al., 2015). Such findings suggest that the HRL effect could still promote long-term memory mechanisms despite interruptions, and is robust to immediate distraction.

5. Conclusion

The current study found that adolescents with non-specific ID showed HRL effects that were similar in magnitude to children with TD matched for mental age, supporting a developmental approach to ID (albeit with some uncertainty given small overlaps in credible intervals across groups). The findings suggest that for individuals with ID, repetition of sequences within immediate serial recall tasks improves long-term serial order learning via the gradual integration of serial order information from short-term memory into a more stable long-term memory trace. HRL was found for both verbal and visuospatial materials, supporting the suggestion that it is a domain general process; although contrary to expectations, for our measure of HRL, the effects were larger for verbal than visuospatial materials. The hypothesised links between HRL and vocabulary acquisition as well as other processes, suggest that using repetition in educational contexts could support learning in children and adolescents with ID.

What this paper adds

Hebb repetition learning has been linked to important developmental and educationally relevant abilities including vocabulary, grammar and reading.

Hebb repetition learning has not previously been studied in adolescents with non-specific intellectual disabilities: we provide information about this process.

Hebb repetition learning occurred in adolescents with intellectual disabilities and its magnitude was broadly similar to that of younger mental age matched children with typical development.

For both groups, Hebb learning occurred in verbal and visuospatial modalities, thereby supporting the argument that Hebb repetition learning is a domain general process.

Our findings about adolescents with intellectual disabilities suggest there could be benefits of using this form of learning to support the acquisition and strengthening of their language and literacy abilities in educational contexts.

CRediT authorship contribution statement

Lucy Henry: Conceptualisation, Methodology, Validation, Resources, Data curation, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **David Messer**: Conceptualisation, Methodology, Validation, Writing – review & editing, Funding acquisition. **Sebastian Poloczek**: Conceptualisation, Methodology, Validation, Data curation,

Declaration of Competing Interest

None.

Data Availability

On publication, the original data and analysis files will be uploaded to the Open Science Framework: https://osf.io/gkpwh/.

Acknowledgements and Funding

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ridd.2022.104219.

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Association.

Archibald, L. M. D., & Joanisse, M. F. (2013). Domain-specific and domain-general constraints on word learning. *Memory & Cognition, 41*, 268–280. https://doi.org/10.3758/s13421-012-0259-4

Attout, L., Ordonez Magro, L., Szmalec, A., & Majerus, S. (2020). The developmental neural substrates of Hebb repetition learning and their link with reading ability. *Human Brain Mapping, 41*(14), 3956–3969. https://doi.org/10.1002/hbm.25099

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Bogaerts, L., Siegelman, N., Ben-Porat, T., & Frost, R. (2018). Is the Hebb repetition task a reliable measure of individual difference in sequence learning? *Quarterly Journal of Experimental Psychology, 71*(4), 892–905.

Bogaerts, L., Szmalec, A., De Maeyer, M., 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

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

Burack, J. A., Evans, D. W., Russo, N., Napolean, J.-S., Goldman, K. J., & Iarocci, G. (2021). Developmental perspectives on the study of persons with intellectual disability. *Annual Review of Clinical Psychology, 17*, 13.1–13.25. https://doi.org/10.1146/annurev-clinpsy-081219-090532

Charlton, C., Rasbash, J., Browne, W. J., Healy, M., & Cameron, B. (2018). *MLwiN version 3.02. Centre for Multilevel Modelling.* United Kingdom: University of Bristol.

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34*(8), 1720–1729.

Hebb, D. O. (1961). Distinctive Features of Learning in the Higher Animal. In J. F. Delafresnaye (Ed.), *Brain mechanisms and learning* (pp. 37–46). Oxford: Blackwell.

Henry, L. A., Messer, D. J., & Poloczek, S. (2018). Working memory and intellectual disabilities. Ch. 2. In T. Alloway (Ed.), *Working Memory and Clinical Developmental Disorders: Theories, Debates and Interventions* (pp. 9–21). Taylor and Francis. Ch. 2.

Hitch, G. J., Flude, B., & Burgess, N. (2009). Slave to the rhythm: Experimental tests of a model for verbal short-term memory and long-term sequence learning. *Journal of Memory and Language, 61*, 97–111. https://doi.org/10.1016/j.jml.2009.02.004

Hsu, J. H., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science, 17*(3), 352–365.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: an overview of the literature and directions for future research. *Psychological Bulletin, 140*(2), 339–373. https://doi.org/10.1037/a0034221

Jarrold, C., Purser, H. R. M., & Brock, J. (2006). Short-term memory in Down syndrome. In T. P. Alloway, & S. E. Gathercole (Eds.), *Working memory and neurodevelopmental disorders* (pp. 239–266). Hove: Psychology Press.

Johnson, A. J., Dygacz, A., & Miles, C. (2017). Hebb repetition effects for non-verbal visual sequences: determinants of sequence acquisition. *Memory, 25*(9), 1279–1293. https://doi.org/10.1080/09658211.2017.1293692

Kalm, K., & Norris, D. (2016). Recall is not necessary for verbal sequence learning. *Memory & Cognition, 44*, 104–113. https://doi.org/10.3758/s13421-015-0544-0

Lifshitz, H., Kilberg, E., & Vakil, E. (2016). Working memory studies among individuals with intellectual disability: an integrative research review. *Research in Developmental Disabilities, 59*, 147–165. https://doi.org/10.1016/j.ridd.2016.08.001

Lifshitz-Vahav, H., & Vakil, E. (2014). Taxonomy of moderators that govern explicit memory in individuals with intellectual disability: Integrative research review. *Journal of Applied Research in Memory and Cognition, 3*(2), 101–119. https://doi.org/10.1016/j.jarmac.2014.03.001

Logie, R. H., Saito, S., Morita, A., Varma, S., & Norris, D. (2016). Recalling visual serial order for verbal sequences. *Memory & Cognition, 44*, 590–607. https://doi.org/10.3758/s13421-015-0580-9

MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods, 7*(1), 19–40. https://doi.org/10.1037/1082-989X.7.1.19

McKenzie, K., Milton, M., Smith, G., & Ouellette-Kuntz, H. (2016). Systematic review of the prevalence and incidence of intellectual disabilities: current trends and issues. *Current Developmental Disorders Reports, 3*, 104–115. https://doi.org/10.1007/s40474-016-0085-7

Mosse, E. K., & Jarrold, C. (2008). Hebb learning, verbal short-term memory, and the acquisition of phonological forms in children. *Quarterley Journal of Experimental Psychology, 61*, 505–514.

Mosse, E. K., & Jarrold, C. (2010). Searching for the hebb effect in down syndrome: evidence for a dissociation between verbal short-term memory and domain-general learning of serial order. *Journal of Intellectual Disability Research, 54*(4), 295–307. https://doi.org/10.1111/j.1365-2788.2010.01257.x

Oberauer, K., Jones, T., & Lewandowsky, S. (2015). The Hebb repetition effect in simple and complex memory span. *Memory & Cognition, 43*, 852–865. https://link.springer.com/article/10.3758/s13421-015-0512-8.

Page, M. P. A., Cumming, N., Norris, D., Hitch, G. J., & McNeil, A. M. (2006). Repetition learning in the immediate serial recall of visual and auditory materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 716–733. https://doi.org/10.1037/0278-7393.32.4.716

Page, M. P. A., & Norris, D. (2009). A model linking immediate serial recall, the Hebb repetition effect and the learning of phonological word forms. *Philosophical Transactions of the Royal Society of London, 364*, 3737–3753. https://doi.org/10.1098/rstb.2009.0173

Patel, D. R., Cabral, M. D., Ho, A., & Merrick, J. (2020). A clinical primer on intellectual disability. *Translational Pediatrics, 9*(Suppl 1), S23–S35. https://doi.org/10.21037/tp.2020.02.02

Pickering, S., & Gathercole, S. E. (2001). *Working Memory Test Battery for Children (WMTB-C)*. London, UK: Pearson Assessment.

Reynolds, C. R., & Voress, J. K. (2007). *Test of Memory and Learning* (2nd ed.). Austin, TX: Pro-Ed.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales* (5th ed.). Riverside Publishing.

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

Smalle, E. H. M., Bogaerts, L., Simonis, M., Duyck, W., Page, M. P. A., Edwards, M., & Szmalec, A. (2016). Can chunk size differences explain developmental changes in lexical learning? *Frontiers in Psychology, 6*, 1925. https://doi.org/10.3389/fpsyg.2015.01925

Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M., & Szmalec, A. (2018). Children retain implicitly learned phonological sequences better than adults: a longitudinal study. *Developmental Science, 21*(5). https://doi.org/10.1111/desc.12634

Sparrow, S. S., Cicchetti, D. V., & Saulnier, C. A. (2016). *Vineland Adaptive Behavior Scales* (3rd ed.). Bloomington, MN: Pearson,.

Sukegawa, M., Ueda, Y., & Saito, S. (2019). The effects of Hebb repetition learning and temporal grouping in immediate serial recall of spatial location. *Memory & Cognition, 47*, 643–657. https://doi.org/10.3758/s13421-019-00921-9

Weiss, B., Weisz, J. R., & Bromfield, R. (1986). Performance of retarded and nonretarded persons on information-processing tasks: further tests of the similar structure hypothesis. *Psychological Bulletin, 100*(2), 157–175.

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), Article e12552. https://doi.org/10.1111/desc.12552

Yanaoka, K., Nakayama, M., Jarrold, C., & Saito, S. (2019). Determining the developmental requirements for Hebb repetition learning in young children: grouping, short-term memory, and their interaction. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(4), 573–590. https://doi.org/10.1037/xlm0000606

Zhang, Z., Parker, R. M. A., Charlton, C. M. J., Leckie, G., & Browne, W. J. (2016). R2MLwiN: A package to run MLwiN from within R. *Journal of Statistical Software, 72*(10), 1–43. https://doi.org/10.18637/jss.v072.i10