UNIVERSITYOF BIRMINGHAM

Building a Lego wall: Sequential action selection

Arnold, Amy; Wing, Alan; Rotshtein, Pia

Document Version
Peer reviewed version

[Link to publication on Research at Birmingham portal](#)

Building a Lego Wall: Sequential Action Selection

Amy Arnold*, Alan M. Wing and Pia Rotshtein*

SyMoN Lab, School of Psychology, University of Birmingham.

Birmingham. B15 2TT. UK

* Corresponding author: Amy Arnold email: arnldamy@gmail.com

*Pia Rotshtein, email: p.rotshtein@bham.ac.uk

Abstract

The present study draws together two distinct lines of enquiry into the selection and control of sequential action: motor sequence production and action selection in everyday tasks. Participants were asked to build two different Lego walls. The walls were designed to have hierarchical structures with shared and dissociated colors and spatial components. Participants built one wall at a time, under low and high load cognitive states. Selection times for correctly completed trials were measured using 3D motion tracking. The paradigm enabled precise measurement of the timing of actions, whilst using real objects to create an end product. The experiment demonstrated that action selection was slowed at decision boundary points, relative to boundaries where no between-wall decision was required. Decision points also affected selection time prior to the actual selection window. Dual task conditions increased selection errors. Errors mostly occurred at the boundary between chunks and especially when these required decisions. The data support hierarchical control of sequenced behavior.

Public significance

Activity of daily living (ADL) can be described as a sequence of actions that are hierarchically organized. For example, the task of making an instant coffee, can be broken down to more specific sequence of actions: boiling the kettle, adding coffee to the mug, pouring boiling water etc... Apraxia and action disorganization are neurological syndromes (AADS) associated with failure to execute sequenced actions, leading to loss of independence.

In the context of laboratory experiment, the mechanism underlying learning and execution of sequenced actions has been investigated using simple key pressing tasks. There are many differences between action sequence in the context of ADL and key pressing tasks. The present study aimed to bridge between the two research fields. We designed a novel Lego building task that comes midway between laboratory key pressing and ADL tasks. Our main aim was to identify weak points in the sequence that are susceptible for errors. Participants were asked to build two partially similar Lego walls. We found that error and slow responses occurred at the dissimilarity points between the walls. This suggest that rehabilitation of patients who suffer from AADS should account for overlap in action units across ADL tasks.

Introduction

Processes underlying automated action sequences have been examined in various daily activity and laboratory settings. This paper aimed to form a bridge between everyday actions and lab based key pressing tasks, by introducing a new sequencing task that falls midway between the two.

Sequential behavior is often described using a hierarchical structure, made up of increasingly complex subunits (Botvinick, 2008; Lashley, 1951; Schneider & Logan, 2006; Schwartz, 2006). Both Motor Sequence Production, with rapidly produced small amplitude movements and everyday tasks, such as washing and dressing can be described in terms of hierarchical structure. Whether sequences are hierarchically processed is less clear. Alternative views suggest that learned sequences are executed through lateral associations (Logan & Crump, 2011; Botvinic and Plaut, 2004) or chunking of responses (Varwey and Abrahamse 2012).

Frameworks across all views make a broad distinction between higher level *selection* processes engaged in retrieving and planning sequences based on symbolic representations, and lower level *execution* processes involved in muscle synergies and posture-based motor planning (Rosenbaum, Cohen, & Dawson, 2009). The interaction between the processes shifts according to task demands and level of task automatization.

For activities of daily living (ADL), a dual system model (Norman & Shallice, 1986) is proposed to explain the sequenced behavior. The model postulates a supervisory system which controls and monitors a contention scheduling (CS) system. The CS was inspired by the schema theory proposed for discrete skill motor sequence (Schmidt, 1975). According to this model (Cooper and Shallice, 2006), routine tasks are executed through an interplay between: 1) hierarchical representations of action schemas, 2) representations of objects in the environment and 3) availability of effectors (e.g. hands). Each action schema is associated with a clear goal and is activated via external triggers (objects) or an internal (higher order) "source" schema. A 'Boiling kettle' schema can be

activated by seeing a kettle or through a higher schema of making tea. Reduced top down control is evident if external triggers have a larger impact, hindering the achievements of goals.

A schema also specifies the flexibility of executing lower schemas. The flexibility is reduced when moving down the hierarchy. For example, the order in which "add milk" or "add sugar" schemas are used is flexible, but once "add milk" schema is selected it triggers a fixed order of lower sub-goal schemas (e.g. pick milk jar [reach -> grasp -> lift]). Schemas are in lateral inhibitory competition (Cooper and Schalice, 2006). Selection of schema would be slower and prone to errors at a flexible relative to lower inflexible layers. This prediction resonates with observed errors reported in healthy participants and neurological patients, where errors of omitting or adding entire sub-goal schemas are more common than errors within lower schemas (Reason, 1979, 1984; Rumelhart & Norman, 1981; Schultz et al., 1991;   Bickerton, Humphreys, & Riddoch, 2006; Forde & Humphreys, 2000; Morady & Humphreys, 2008; Schwartz et al., 1999). The dual system model, like many other models of ADL, is silent in relation to the execution phase. This may be because movement associated with ADL is complex and varies substantially depending on the objects (e.g. electrical kettles vary in the specific movement required to operate them).

A Lab based simplified versions of well learned sequence task have been developed to tap into the execution phase. For example, the Discrete Sequence Production (DSP) tasks require participants to execute explicitly learned sequences as fast as possible – either following external cues (Abrahamse, Ruitenberg, de Kleine, & Verwey, 2013) or from memory (Wiestler & Diedrichsen, 2013). These tasks often rely on key presses, but some have used more elaborate movements (e.g. Panzer, Krueger, Muehlbauer, Kovacs & Shea, 2009; Shea, Kovacs & Panzer, 2011). We focus on the DSP task as it is understood to be representative of more complex real-world action (Abrahamse et al., 2013). Performances in these tasks are typically measured as the interval between two consecutive responses, or the time to initiate the first response.

The Dual Processor Model (DPM, Verwey, 1999, Abrhamase, 2013) has been proposed to explain various observation arising from the DSP. In contrast to the ADL model introduced above, the DPM does not assume hierarchical structure, in which higher level cognitive representations control/guide lower level motor-based representations. Instead it suggests sequenced movement executed by cognitive and motor processors. These two processors operate in parallel and provide partially redundant information (Verwey 2001, Abrahamse, 2013). The cognitive processor uploads actions one-by-one to a motor buffer, enabling production of complex and novel sequence actions (Abrahams, 2013).

Learning is reflected by the grouping of actions to chunks. This enables both the cognitive processor and the motor processor, to process multiple actions as one unit. This chunking facilitates production by reducing the time it takes to upload and retrieve information from the motor buffer (Wymbs, Bassett, Mucha, Porter, & Grafton, 2012).

Typically, with longer sequences (6-7 keys), chunks emerge spontaneously and their size varies between participants as a function of individual's working memory (Bo, Borza, & Seidler, 2009; Miller, 1956; Sakai, Kitaguchi, & Hikosaka, 2003). Chunk structure can be controlled through introducing pauses (Verwey, Abrahamse, & de Kleine, 2010) and color grouping (Jimenez & Vázquez, 2011). Effects of sequence selection are evident typically at the beginning of the sequence, when the programing of the entire sequence is uploaded to working memory. Initiation of the first response of a chunk, within a sequence is also slowed. This reflects a concatenation point, the time to upload the chunk into the motor buffer (see Abrahamse 2013 for review). Selecting and uploading of the chunks is done by the cognitive processor. Once a chunk is uploaded its execution is automatic, as it is not slowed by a dual task (Verwey, Abrahamse, & de Kleine, 2010; Verwey, Abrahamse, De Kleine, & Ruitenberg, 2014). However, the cognitive processor is involved at concatenation points. Though it is assumed to be a low demanding process, as the cognitive processor can in parallel support performances in other tasks, with graded latency costs as function of task difficulties (Verwey, Abrahamse, De Kleine, & Ruitenberg, 2014).

In a classic DSP task overlaps between sequence structures is not controlled for (e.g. Verwey et al., 2001). Rosenbaum and colleagues (Rosenbaum, Inhoff & Gordon, 1984) examine the impact of 'competition' using partially overlapping sequences. Partial dissimilarity slows the initiation of the first response, potentially reflecting a cost of resolving the upcoming conflict (Though see Rose, 1988). Like for the ADL tasks, Rosenbaum and colleagues (1984) suggest that slowness of the responses echoes the selection processes between lateral segments of hierarchical structured sequence. The description units are grouped based on similarity and dissimilarity across sequences (Rosenbaum, Cohen, Jax, Weiss & van der Wel, 2007).

In summary, "weak" points in the sequence, measured by slowing of response and increase errors are attributed to two sources: 1) a selection process between competing lateral sub-actions or, 2) uploading of the next motor chunk. In this study we aimed to test whether sequences are processed hierarchically by assessing the impact of uploading and selection processes at chunk boundaries using a task that comes midway between DSP and ADL.

Participants were trained to build two hierarchically structured walls using Duplo Lego bricks (Figure 1). This enabled the investigation of sequential performance using real objects in a controlled setting. Chunk structure of the Lego walls was created using colors as well as pauses in cue presentations. The task delineated structural boundaries between chunks where a *decision* between the two walls was required and boundaries where no decision needed to be made. All other points of measurement were *within chunk* actions. Performances were recorded with motion tracking and video cameras, enabling the exact coding of errors and inter-action intervals (IAI). We asked whether errors and IAI differed at boundary and decision points relative to within chunks. To assess a potential role of a generic cognitive processor we also measured performances under a dual task condition.

Method

Detailed description of the methods is presented in Supplementary materials.

Participants

10 participants (mean age = 24.3yr) took part in two experimental sessions, separated by one day. Each session lasted ~ 1.5 hours. All participants were right handed. The study was approved by the local ethical committee.

Procedure and Tasks

Participants were asked to build two Lego walls (Anna's wall and Daniel's wall – Figure 1). The experiment was divided into four blocks. Day 1 included two training blocks (1 & 2). Day 2 included two test blocks (3 & 4). Each block had 20 trials, 10 per wall. Block 4 was performed under dual-task condition, in which the secondary task involved monitoring an audio sequence for the highest uttered number (Ruh, Cooper, & Mareschal, 2010).
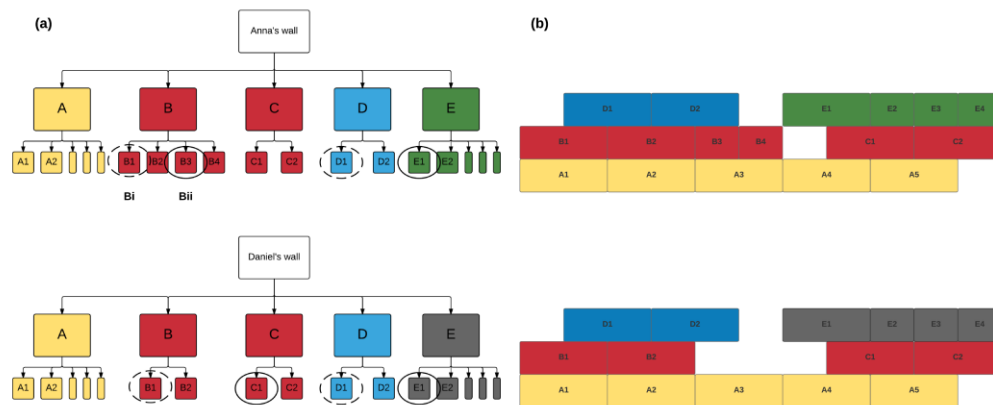
Figure 1. The Lego wall
a) Hierarchical representation of the task walls. Task goals are the wall names (Anna and Daniels). Each wall consists of chunks (A=yellow, B-right=red, C-left=red, D=blue, E=green/black) and each chunk is made up of individual bricks (e.g., from right-to-left: E1 to E4). Solid ellipses mark decision boundary points (dissimilar chunks across walls) and dashed ellipses mark no decision boundary points (identical chunks walls). **b)** Schematic representation of the complete walls structure. Labelling does not necessarily represent the order of brick placement.

During the experiment, participants sat at a table with two Duplo Lego baseplates in front of them. The one on the right contained Lego bricks (Figure 1) used to build the wall on the left baseplate.

At the beginning of each trial, participants were presented with a picture of the target wall and started building the wall only after it disappeared. To ensure no systematic errors, the target wall was shown again as feedback at the end of the trial. We recorded error and timing by video and motion tracking cameras.

Data analyses

We used the video data to compute overall completion time (reported in Supp. Material) and record errors.

Selection time (inter action interval, IAI) was defined as the time placing a brick on the wall, till picking a new brick from the right baseplate (sFigure 3). IAIs were calculated for correctly completed trials only. Bricks from the first layer ('A') were not analyzed. We defined four bricks of interest based on their position in a chunk: 1st brick (boundary) vs. 2nd or later (within chunk) and depending on the structural overlap between the two walls: non-overlapping chunks, requiring a decision or overlapping chunk, requiring no decision.

Further analyses considered only the bricks at the 2nd position, to test whether *within chunk* bricks should be further divided to *pre-decision* points (the brick before a *decision boundary*) or *predict* points (the brick before a no decision boundary, i.e. structurally predictable point).

Results

Participants learned to build the two walls, and after 5 practices of building each wall errors on completion were minimal. Furthermore, speed of total completion time stabilized from the second practice block onward to less than 30sec per wall. Participants were also accurate at completing the secondary auditory task with an average of 88.5% accuracy. Though, the dual task did not affect overall completion time. For more detail see Supp. Material.

Error analysis

Out of 400 testing trials, 98 (range 6-16 across participants) included at least one error, and only one was not corrected before completion. Eighty-four were cognitive errors and fourteen handling errors (Table 1). A 2 (nDT, DT) x 2 (boundary type: decision, no decision) factorial design was used to compute differences in the frequency of the cognitive errors. A reliable difference between conditions is observed, $X^2(3)=24.5$, $p<0.001$, d=1.33. Simple comparisons showed that the dual task (relative to nDT) increased errors mostly at the *decision boundaries*, $X^2(1) =9.62$, $p=0.002$, d=0.92; but did not affect errors at *no decision points*. During the dual task, there were more errors at decision than no decision points $X^2(1)=13$, $p<0.001$, d=1.34. Though, in both blocks there were more errors at boundary than within chunk points, nDT: 28 vs. 2, $X^2(1)=22.5$, $p<0.001$, d=3.46; DT: 42 vs. 2, $X^2(1)=46.3$, $p<0.001$, d>6.

Examining the type of errors made, showed that the dual task affected mostly brick's *selection* (nDT=7 vs. DT=26) rather than *placement* (nDT=9 vs. DT=9). A two (Block) by two (error type) Chi Square test confirmed reliable difference in the distribution of error types across blocks, $X^2(3)=18.57$, p<0.001, d=3.25. With significant larger number of selection than placement errors in the dual task block, $X^2(1)=10.94$, P<0.005, d=1.35; and also larger number of selection errors in the dual task than non-dual task block, $X^2(1)=8.26$, P<0.005, d=1.5.

There were also more handling errors in the dual task than the no dual task

block. The data suggests that the dual task increased the number of errors, especially at decision points, primarily interrupting the selection processes.

Table 1

Location and frequency of self-corrected errors.

|	No Dual Task	Dual Task	Total
Decision boundary points	16	39	**45**
No decision boundary points	12	13	**22**
Within chunk points	2	2	**4**
Handling	4	10	**14**
Total	**34**	**64**	

Selection latencies analysis (Inter-action intervals) within the testing blocks

Data of 2 participants were excluded. These participants did not have a consistent sequence of building the wall throughout the experiment (< 16 trials with the same order). This meant that we could not define decision vs. no decision boundary points for these participants.

A 2 (chunk: *decision/no decision*) x 2 (brick position: *boundary/within chunk*) x 2 (task: *dual task/no dual task*) repeated measures ANOVA was used to analyze the IAI data. The dual task did not affect the IAI, nor interacted with any of the factors. There was a *chunk type x brick position* interaction, $F(1,7)=47.97$, $p<0.001$, d=3.46 (Figure 2a). Paired-samples t-tests showed that in *decision* chunks, IAIs were longer at *boundary points* than *within chunk* points, $t(7)=4.78$, $p=0.002$, d=2.39. Surprisingly, in *no decision* chunks the pattern reversed, $t(7)=3.34$, $p=0.012$, d=1.67. This was an unexpected effect.

One explanation for the above reverse effect is that in the context of the current set-up, *within chunk* bricks can also function as *pre-decision* points (for example,

D2 precedes E1; Figure. 2a). Therefore, we next tested whether *within chunk* bricks that reliably served as*, pre-decision* points took longer to select than *within chunk* bricks that served as *predict* points.
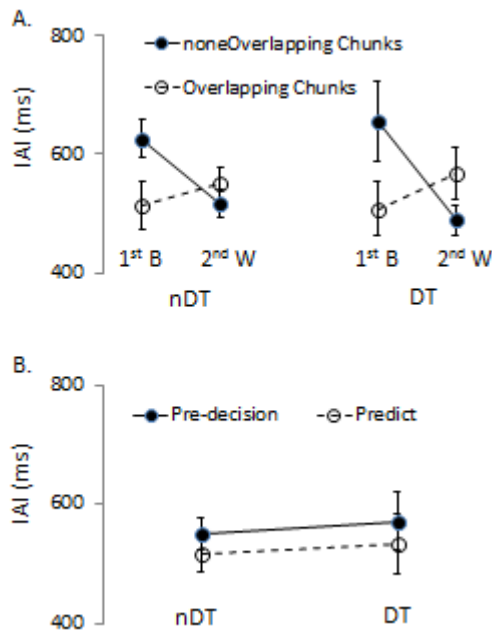
Figure 2. Inter Action Interval
A. the plot represents the IAI (selection of a brick) at different functional points. Solid line bricks from none overlapping chunks that require decision; Dashed lines bricks from overlapping chunk with no decision. 1st B, selection of 1st brick at a boundary; 2nd W, selection of within chunk's brick;
B. IAI for within chunk bricks, solid line pre-decision bricks, preceding none overlapping chunk; dashed line predict bricks, preceding overlapping chunks. nDT, no dual task block; DT, dual task block.

A 2 (function: *pre-decision/predict*) x 2 (task: *DT/nDT*) repeated measures ANOVA revealed a main effect of *function*, with selection times of within chunk bricks at *pre-decision* points taking significantly longer than selection at *predict* points, $F(1,7)=6.68$, $p=0.035$, d=1.29 (Figure 2b). Task did not affect the results or interacted with the condition.

Discussion

In this study we used a novel Lego building task composed of colored bricks chunks to assess execution of sequenced actions. Participants learned to build two partially overlapping structured walls. We focused on selection processes underlying transition between chunk boundaries. To this aim, we defined two types of boundaries: 1) decision boundary reflecting a transition to a dissimilar chunk across sequences, and 2) no decision boundary between overlapping

chunks across the two walls. We observed more cognitive errors at chunk boundaries than within chunks. More interestingly, we found increased errors and slower inter action interval (IAI) when selecting bricks at a decision relative to no decision boundary. We also observed that IAI slowed when selecting for the action preceding a decision relative to an action that did not precede a decision. Surprisingly, 'no-decision chunk' boundary did not slow responses compared to within chunk responses. A secondary aim examined whether selection processes were affected by a dual task. The number of errors increased in the dual task condition, specifically at the decision boundary. These were primarily selection errors rather than placement errors. The dual task did not affect IAIs. Taken together, the data support hierarchical models for routinized sequence tasks.

The observation that selection process at decision boundary points were most vulnerable to errors (especially under dual task condition) and costly in terms of timing, is in line with models postulating that sequenced actions are executed using hierarchical structure (Cooper and Shallice, 2006, Rosenbaum et al., 2007, Rhodes et al., 2004). It is assumed that at these points competition between alternative chunks, required traversing the hierarchy 'to consult' higher order schemas. While a simple chunk boundary involving no decision were potentially executed through lateral association triggers and affected selection only marginally through increase errors.

A further key finding is that preparation for upcoming decision points affected selection times of the immediately preceding *within chunk* bricks, indicating that preparation for a difficult point in the task begins during the processing of previous actions, before the selection of the action itself. This is in line with Rosenbaum and colleagues' observations (Rosenbaum et al., 1983, 1984) suggesting that dissimilarity between sequences affect preceding responses (see also Supp. Results). Similarly, in typing, finger movements for one keystroke begin before finger movement for the preceding keystroke ends (Flanders & Soechting, 1992). The notion that cognitive processes can occur in the background and prior to the required action has also been shown for classic DSP

tasks, where a cost of Simon conflict on response was diminished when preceded by other non-conflicting responses (Verwey, 1995).

Surprisingly, we did not observed chunk concatenation effects on the IAI (Verwey, 2001), even though we used colour to elicit chunks' boundaries. When the transition between chunks was identical across both walls (no decision boundary) latencies at concatenation points were no different than within chunk IAI. We note that error at chunk boundaries increased at both decision and no-decision points. It could be that the large amplitude movement concealed any preparation effect (select and upload), without disrupting the temporal rhythm. It is also possible that in the context of two competing sequences, chunks are defined based on similarity and dissimilarity across sequences rather than properties of the stimuli - colour (Rosenbaum, Cohen, Jax, Weiss & van der Wel, 2007).

We used an auditory dual task, to overload the cognitive system. This manipulation led to increase of errors. As predicted by Cooper and Shallice (2006) model, reduced top down control, meant larger impact on the selection from the objects in the environment (than the relevant schema's wall), leading to increased selection errors. Overloading the cognitive system did not affect the time it took to select a brick or the time it took to complete the wall. It is difficult to account for the lack of timing effect here. It could be that this effect was masked by noise arising from the complex movement used here. It could also be that other more demanding tasks will show larger interference (Verwey et al., 2014).

We designed the Lego Walls paradigm as a bridge between lab-based simple key-press and ADL tasks. A few limitations should be noted as a precursor to any conclusions. First, the sample size of the current study was relatively small (n=10) and was composed of psychology students limiting the generalizability of the results. Furthermore, 20% of the participants used a flexible rather than more rigid and fixed strategies when constructing the walls. This suggests that underlying mechanisms to execute sequenced behavior may vary between individuals. Finally, in contrast to key-press DSP task which uses hundreds of

training trials, or ADL tasks which uses less than a dozen training trials; our tasks, were designed to be in between, had twenty training trials. As training is suggested to introduce a qualitative change to the way actions are executed, any direct comparisons between the results of the current study and past findings should be made with caution. Nevertheless, we believe that our results demonstrate that when using partially overlapping sequences, points of divergence that require decision take longer to execute, and under dual task condition increase selection errors. We also showed that preparation starts in the background of preceding actions, slowing their executions. Together, our data support hierarchical processing in the execution of sequential behavior.

The current task provides a methodologically novel approach to investigate the performance of complex sequential tasks. Collecting timing data using real objects is time intensive, making it difficult to reproduce the practice levels of motor sequence tasks and other highly skilled tasks, such as typing and musical performance. Analyzing movement traces is also difficult to automate and divergence from expected movement can result in lost trials. The current paradigm is unique, as allows precise control over task structure, timing and accuracy, whilst retaining many of the important features of everyday tasks. The reproduction of standard results showing hierarchical control of action suggests continuity between everyday tasks and motor control. We hope that future designs could use the present paradigm to further investigate the link between motor sequence production and everyday action.

References

Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E., & Verwey, W. B. (2013).

Control of automated behavior: insights from the discrete sequence production task. *Frontiers in Human Neuroscience,* 7(March), 82. doi: 10.3389/fnhum.2013.00082.

Bickerton, W. L., Humphreys, G. W., & Riddoch, J. M. (2006). The use of memorised verbal scripts in the rehabilitation of action disorganisation syndrome. *Neuropsychological Rehabilitation*, 16(2), 155–77.

Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory correlate with deficits in explicit motor sequence learning. *Journal of Neurophysiology*, 102(5), 2744–54.

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, 12(5), 201–8.

Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review*, 12(6), 1011–7.

Cooper, R.P. & Shallice T. (2006), Hierarchical schemas and goals in the control of sequential behavior. *Psychological Review*, 113(4), 887-916.

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity Subserving Motor Skill Learning. *Neuron*, 72(3), 443–54.

Flanders, M., & Soechting, J. F. (1992). Kinematics of typing: parallel control of the two hands. *Journal of Neurophysiology*, 67(5), 1264–74.

Forde, E. M., & Humphreys, G. W. (2000). The role of semantic knowledge and working memory in everyday tasks. *Brain and Cognition*, 44(2), 214–52.

Gordon, P. C., & Meyer, D. E. (1984). Perceptual-motor processing of phonetic features in speech. *Journal of Experimental Psychology: Human Perception and Performance,* 10(2), 153.

Jiménez, L., & Vázquez, G. a. (2011). Implicit sequence learning and contextual cueing do not compete for central cognitive resources. *Journal of Experimental Psychology. Human Perception and Performance*, 37(1), 222–35.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior*, (pp. 112-36). New York: Wiley.

Logan, G. D., & Crump, M. J. (2011). Hierarchical Control of Cognitive Processes: The Case for Skilled Typewriting. In B Ross (Ed.), *Psychology of Learning and Motivation-Advances in Research and Theory*, 54, 1, 1-27.

Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2), 81-97.

Morady, K., & Humphreys, G. W. (2008). Comparing action disorganization syndrome and dual-task load on normal performance in everyday action tasks. *Neurocase*, 15(1), 1–12.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and Automatic Control of behavior (pp. 1-18). In R.J. Davidson, G.E. Schwartz & D Shapiro (Ed.), *Consciousness and Self-Regulation Advances in Research and Theory*, Volume 4, Springer, US.

Panzer, S., Krueger, M., Muehlbauer, T., Kovacs & A.J., Shea, C.H. (2009) Inter-manual transfer and practice: coding of simple motor sequences. *Acta Psychologia*, 131(2), 99-109. doi: 10.1016/j.actpsy.2009.03.004.

Reason, J. T. (1979). Actions not as planned: The price *of automatization. In G. Underwood & R. Stevens (eds.), Aspects of Consciousness*. Academic Press 1-67.

Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & R. Davies (Eds.), *Varieties of attention*. New York: Academic Press. 515-549.

Rhodes, B. J., Bullock, D., Verwey, W. B., Averbeck, B. B., & Page, M. P. A. (2004). Learning and production of movement sequences: behavioral, neurophysiological, and modeling perspectives. *Human Movement Science*, 23(5), 699–746.

Rose, D. J. (1988). Choosing Between-Movement Sequences: Effect of Response-Choice Similarity on the Underlying Programming Operations. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4), 638-

645.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology. Human Perception and Performance*, 9(1), 86–102.

Rosenbaum, D. A., Inhoff, D. A., & Gordon, A. M. (1984). Choosing Between Movement Sequences: A Hierarchical Editor Model. *Journal of Experimental Psychology: General* , 113(3), 372-393.

Rosenbaum, D.A, Cohen R. G., Jax S. A., Weiss, D. J, van der Wel R. (2007). The problem of serial order in behavior: Lashley's legacy. *Humam Movement Science*, 26(4):525-54.

Rosenbaum, D. A., Cohen, R. G., Dawson, A. M., Jax, S. A., Meulenbroek, R. G., van der Wel, R., & Vaughan, J. (2009). The posture-based motion planning framework: new findings related to object manipulation, moving around obstacles, moving in three spatial dimensions, and haptic tracking. *Advanced in Experiment Medical Biology,* 629, 485-97.

Ruh, N., Cooper, R. P., & Mareschal, D. (2010). Action selection in complex routinized sequential behaviors. *Journal of Experimental Psychology. Human Perception and Performance*, *36*(4), 955–75.

Rumelhart, D. E., & Norman, D. A. (1981). Analogical processes in learning. *Cognitive skills and their acquisition*, 335-359.

Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research*, *152*(2), 229–242.

Schmidt, Richard A. (1975) A schema theory of discrete motor skill learning.

Psychological Review, 82(4), 225-260.

Schneider, D. W., & Logan, G. D. (2006). Hierarchical control of cognitive processes: switching tasks in sequences. *Journal of Experimental Psychology. General*, *135*(4), 623–40.

Schwartz, M. F. (2006). The cognitive neuropsychology of everyday action and planning. *Cognitive Neuropsychology*, *23*(1), 202–21.

Schwartz, M. F., Buxbaum, L. J., Montgomery, M. W., Fitzpatrick-DeSalme, E., Hart, T., Ferraro, M., … Coslett, H. B. (1999). Naturalistic action production following right hemisphere stroke. *Neuropsychologia*, *37*(1), 51–66.

Verwey, W. B. (1999). Evidence for a multistage model of practice in a sequential movement task. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1693 - 708.

Verwey, W. B. (2001). Concatenating familiar movement sequences: the versatile cognitive processor. *Acta Psychologica*, *106*(1-2), 69–95.

Verwey, W. B., Abrahamse, E. L., & de Kleine, E. (2010). Cognitive processing in new and practiced discrete keying sequences. *Frontiers in Psychology*, *1*, 32. doi: 10.3389/fpsyg.2010.00032

Verwey, W. B., Abrahamse, E. L., De Kleine, E., & Ruitenberg, M. F. L. (2014). Evidence for graded central processing resources in a sequential movement task. *Psychological Research*, *78*(1), 70–83.

Verwey, W. B., Groen, E. C., & Wright, D. L. (2015). The stuff that motor chunks are made of: Spatial instead of motor representations? *Experimental brain research*, 1-14.

Wiestler, T., & Diedrichsen, J. (2013). Skill learning strengthens cortical representations of motor sequences. *Elife*, *2*, e00801.

Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. a, & Grafton, S. T. (2012). Differential recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron*, *74*(5), 936–46.

Supplementary Materials

Methods:

Participants

Ten participants (3 males, mean age = 24.3yr, range = 19-37yr) were recruited using an opportunistic sample, that included mostly lab members. This was because the task was taxing and the apparatus set up was relatively complex. Hence it required motivated and patience participants. Though, all participants were blind to the specific research questions. The relative small number of participants was also dictated by the laboring analysis procedure which relied on defining at individual level the functional role of each brick, and verifying that each action of picking/placing a brick in the video relates to the identified interval of the motion tracking velocity information (see below for details).

Apparatus

Participants set at a table with two baseplates in front of them. The right baseplate presented the Lego bricks, organized by colors and size in 5 columns (see sFigure 1). There were 26 bricks on the baseplate (more than what was needed for any single wall). The environment, available Duplo Lego, was fixed across the walls and was organized to reduce visual search load, though it contained additional irrelevant Lego pieces. The baseplate on the left of the participant, was empty and contained marks on which the walls should be built.
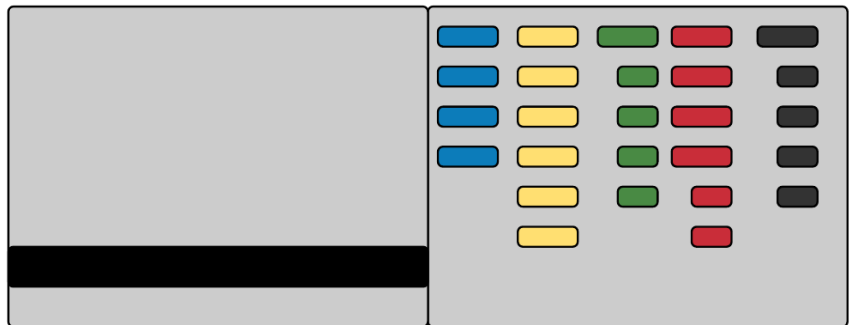

sFigure 1: Experiment set up
The figure presents a schematic representation of the arrangement of the two baseplate. Each colored square represent a Lego brick

The Lego walls

Two Lego walls were used in the experiment. The walls were given a label for ease of reference and memory. The walls consisted of 5 chunks (sFigure 2) defined by the colors and spatial continuity of the bricks. Anna'a wall was constructed of 17 Lego brick and Daniel's of 15 Lego brick. The walls had partly overlapping structures. 'A' (yellow), 'C' (red-right) and 'D' (blue) were identical across both walls. 'B' (red left) and 'E' (black/green) represent two types of non-overlapping structures. Note that change of B also impacted the overall configuration of the two walls and their relations to the other chunks.

The first two units of 'B' were identical across the two walls, we refer to them as 'Bi'; while the second two units appeared only in Anna's wall but not in Daniel's, referred to as 'Bii'. Thus participants had to execute 'Bi+Bii' vs. 'Bi'. This manipulation echoes the manipulation used by Rosenbaum et al., 1984, where responses to partially overlapping sequence structure was manipulated, e.g. an IMR vs. IM.

'E' was the same in terms of its spatial configuration but differed in color between the walls. Anna'a walls used green while Daniel's wall used black Lego bricks. Here ADL hierarchical model predicts slowing when initiating 'E' relative to 'D', due to the need to select the appropriate brick color; DPM also predicts slowing on initiating 'E' relative to 'D', but this is because it is a longer sequence with 4 elements compare to the 2 in 'D'.
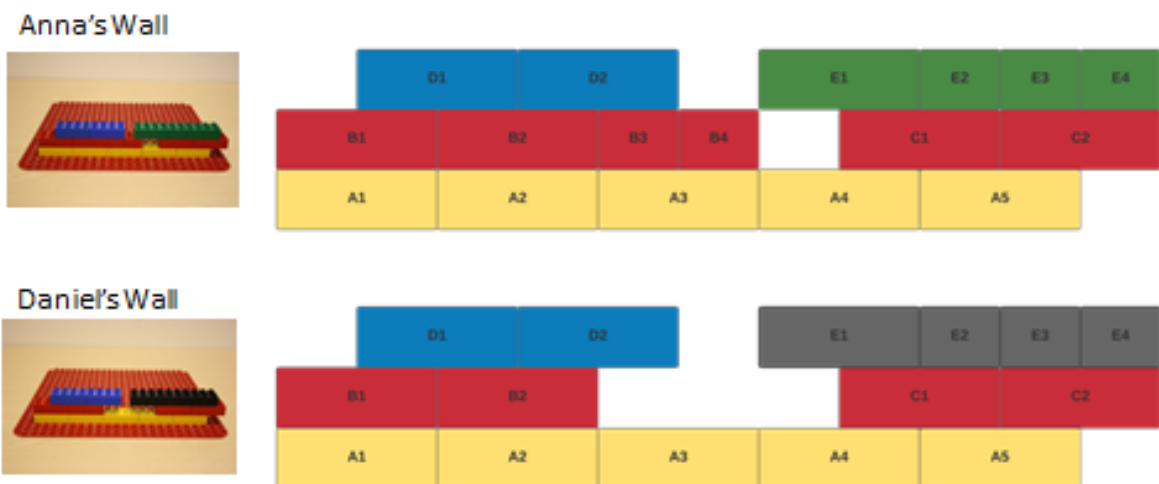
sFigure 2: The structure of the Lego Walls
On the right, the final pictorial cue present to the participants' as the goal. On the left a schematic representation of each wall with labeled colored chunks. The order of the labeling does necessarily correspond to the order of building.

Procedure

At the beginning of each trial, participants were presented with the target wall. To ease comprehension and encouraged chunking, the wall was presented layer-by-layer at 0.3Hz (3sec = Layer 1 (A=yellow bricks); 3sec = Layer 1 + layer 2(B + C red bricks); 3sec completed wall. The presentation order of the walls was random (apart from the first 10 practice trials, where each wall was presented consecutively 5 times). The walls were presented using E-prime 2.0.

Participant started to build the wall once the picture had disappeared. They had unlimited time to complete the wall and were asked to be accurate. At blocks 2-4 they were asked to complete the wall as fast as possible. They were instructed to complete a chunk (same color Lego bricks) before moving to the next, in any order they wished. They could use of both hands when needed, but were also asked to use their right hand to pick and place the Lego bricks. After the building was completed, participants were presented with a photograph of the completed wall as feedback. Feedback was available for 12 seconds in block 1 and for 5 seconds blocks 2-4. This was done to ensure that errors were not systematically structured into the task. Following the feedback, the experimenter dissembled the wall and returned the bricks to the left baseplate, before the next trail commenced.

The dual task

We used an auditory monitoring task similar to that used in (Ruh, Cooper, & Mareschal, 2010). This involved a female voice uttering numbers at an irregular pace with an average rate of 2Hz. The numbers were uttered together with the appearance of the first pictorial cue and ended when the participant completed

the wall. At the end of the trial the participants were asked to report the highest number they heard. The experimenter recorded the correctness of the response. Numbers were randomly drawn from integers between 1 and 60. The upper limit of the range of the distribution was also determined randomly. A secondary auditory task was selected to ensure that effects observed, arises from the engagement of the cognitive supervisor and not due to within modal interference. The block with the dual task was always last, as task familiarity (number of practices) is expected to improve task performances while cognitive load is predicted to impinge on them.

Data collection

We collected video and motion tracking data. Video recordings were made using a LogiTech USB webcam. The camera was placed to enable a clear view of both hands, baseplates and walls. Movement data was collected only on the second day session, blocks 3-4 from the testing. During this session we attached a small reflective marker to the metacarpophalangeal joint of the right hand. Motion of this marker was captured using 4 Qualisys™ 3-D Motion Capture cameras located at the ceiling to reduce occlusions.

Data analyses

We used the video data to compute overall completion time and record errors. Completion times on a trial-by-trial basis were calculated as the mean time (in seconds) to build a wall. This was computed from the time the cue disappeared until the last brick was placed. All trials (correct and incorrect) were included in this analysis. We used one-way ANOVA with 4 levels representing in each block to analyze completion times. This analysis followed by t-tests to specifically assess the effects of training and the dual task on performances.

Error analysis focused only on the testing blocks 3 & 4, completed on day 2. We used X^2 to assess differences between conditions, with error frequencies across participants, as the dependent variable.

The video data in combination with the motion tracking data was used to identify selection time of specific bricks. Selection time (inter action interval, IAI) was defined as the time between letting go after placing the preceding brick on

the wall, till picking a new brick from the right baseplate. This two time points were identified using the movement velocity (sFigure 3). Zero velocity when placing of a brick was defined as the onset, zero velocity when picking a new brick was defined as the offset. Thus in this experiment, IAI reflected the selection process. The IAI could include durations in which the hand hovered above the baseplate, before a decision of which brick to pick was made. IAIs were calculated for correctly completed trials only; trials with no handling errors (e.g., dropping a brick) and no self-correction errors (e.g. changing am initial selection).
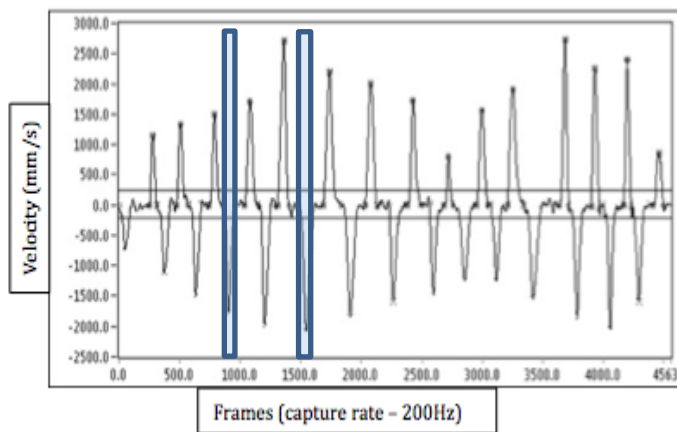
sFigure3: Velocity of a typical trial
The chart presents a velocity profile for a typical trial. Inter Action Intervals were computed between the zero velocity points. Examples are marked with stripes.

Main analysis of IAI as a function of decision and location: We defined four bricks of interest based on their position in a chunk: 1st brick (boundary) vs. 2nd or later (within chunk), and on the structural overlap between the two walls: non-overlapping chunks, requiring a decision while overlapping chunk, requiring no decision. For example, the 1st brick of 'Bii' in Anna's wall (or 'C' in Daniel's Wall) was a decision boundary point between non-overlapping chunks, as participants had to decide whether to continue with B or move to C. The first brick of 'E' was also a decision boundary point, as participants had to decide whether they are using a green or a black color. The 2nd brick in 'Bii' (or 2nd, 3rd and 4th of 'E') were non-boundary points within non-overlapping chunks. The 1st brick of 'Bi' (or 'D') was a no decision boundary of overlapping chunks; while the 2nd was a no decision brick within overlapping chunks.

Importantly, the functional definition of each brick was done individually, based

on the sequences used consistently by the participants. For this reason, participants who had shown no consistent pattern (chunk order) when building the walls were excluded from the IAI based analyses. Bricks from the first layer ('A') were not analyzed. We used repeated measured ANOVA with the following three factors: a 2 (chunk type: decision, non-decision) x 2 (brick position: 1st, $\geq$2nd) x 2 (task: nDT, DT) to analyze the IAI data.

We complemented the main analyses by looking at effects across walls in supplementary analysis.

IAI at Pre-decision versus predict points: Further analyses considered only the bricks at the 2nd position. To test whether *within chunk* bricks should be further divided to *pre-decision* points (the brick before a *decision boundary*) or *predict* points (the brick before a no decision boundary, i.e. structurally predictable point).

Results

Performances in the dual task

Overall participants were highly accurate (88.5% ± 5.8std. range from 80% accuracy to 95%) in completing the secondary task, i.e. reporting the highest number they heard. While numerically auditory memory was better at trials with no errors (92.2 ± 8.3std) than trails with errors (82.3 ± 17.7std) this effect was not reliable. The data clearly show that participants attended the task. As accuracy was not at ceiling it suggests the task was demanding.

Completion time effects across the four blocks

This analysis aimed to establish: 1) whether training improved the performance of the task; 2) whether the performance improvement was stable over the two day sessions; and 3) whether the dual task impinged on performance speed. A relative stabilization of completion times can be seen by block 2


25
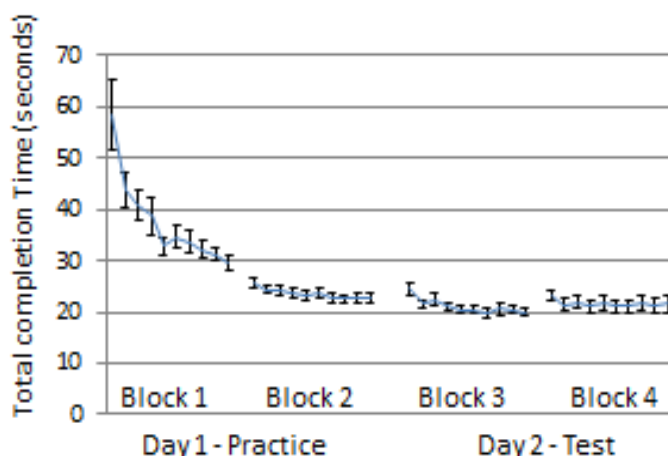
(sFigure 4), which persist in day two. Average completion times in seconds were bl1 = 53.55 ± 10.4std, bl2 = 32.7 ± 3.16, bl3 = 29.58 ± 2.99std, bl4 = 29.86 ± 4.87std. This was verified with an ANOVA, showing reliable effects of block, $F(3, 27) = 70.17$; $p < 0.001$ d = 2.65. Block 1 was the slowest relative to all other blocks, all $t(9) > 7.7$, P < 0.001, d > 3.34. Block 2 was slower than block 3 and 4, $t(9) > 2.55$, P < 0.031, d > 1.14. There were no reliable differences in completion times between blocks 3 and 4, $t(9) = -0.32$, P = 0.75, d=0.14. Thus overall performance, measured as total completion time did not deteriorate following the introduction of the dual task.

sFigure 4: Total completion Time
Averaged time across participants taken to build the walls as a function of practice. Each block had 10 repetitions per wall. All trials, correct and incorrect included in this analysis. Error bars are standard error of the mean.

Supplementary analysis - the wall type as a factor:
We present in sFigure 5, the IAI for all time points in the sequence for each wall at the two testing blocks.
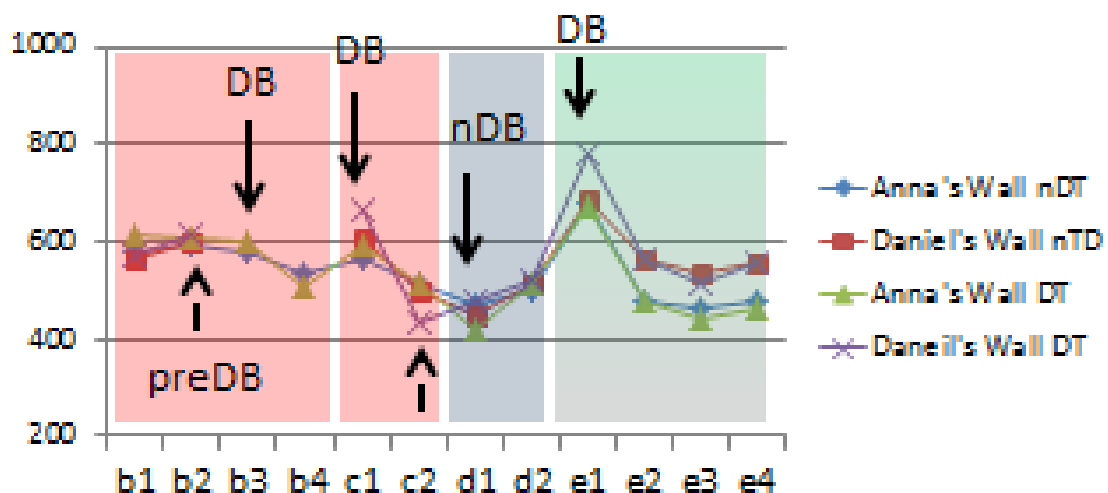
- Evidence for quasi-hierarchical structure:

The 4 elements red bricks chunk (B) was structurally divided to two chunks (Bi, Bii). This was based on the overlap and dissimilarity between the structure of the two walls. This quasi-hierarchical structure of breaking chunks to smaller repetitive units (e.g. *adding one, adding two sugars*) is a feature of many everyday tasks (Botvinick & Plaut, 2004) which is often associated with increase errors (Reason 1979, 1984). We wanted to assess whether the selection IAI represented this quasi-structure with B being broken to separate elements, where Bii.1 is a decision boundary point for Anna's wall as C.1 for Daniel's wall.

Hierarchical based models and the dual processor model (DPM, Verwey 2001, Abrahamse, 2013) make number of opposite predictions on the impact of this manipulation on performances:

1. Hierarchical models (Cooper and Shalice, 2006) predict increase latency and errors at 'Bii' as this is a decision boundary point. But the DPM predict no slowing down at this point, as this is within chunk point (same color as Bi). We computed a 2(wall) x 2(task) x 2(boundary type: D.1, Bii.1/C.1) ANOVA. The results show a main effect for boundary type, $F(1,7) = 36.57$, $p = .001$ $d = 3.02$. IAI at the decision boundary took longer than at no-decision boundary which did not interact with the wall or the task.

2. Hierarchical models (Rosenbaum et al., 2007) predict that initiating 'Bi.1' would be slower relative to initiating 'D'. This is because 'B' is different across the two walls while 'D' is identical. The chunk based model, DPM, predicts that initiation latencies would be affected by the number of units in a chunk. Hence initiating 'D' would be faster only relative to initiating 'B' when building Anna's wall, as the former have 2 units, while the later has 4. When building Daniel's wall time to initiate 'B' and 'D' should not differ. We computed a 2(wall) x 2(task) x 2(boundary type: Bi.1, D.1). The results show a main effect for boundary type,

F(1,7) = 10.11, p = .015 d = 1.58 and no interaction with the wall type, F(1,7)=2.1, p = .19 d = 0.7. Initiating B took longer than D, even though D & B (at least in Anna's wall has the same number of step in the chunk). This support the hierarchical model, replicating Rosenbaum et al., (1984) findings.

- Testing for no decision chunk effects

To assess for evidence of chunking in the data in the absence of decision we compared IAI of 'D.1', initiating a new blue chunk which is identical across the walls to 'E.2', which is within chunk brick. We used again a 2(task) x 2(walls) x 2(bricks) repeated measured ANOVA. We surprisingly observed an interaction of wall and brick, F(1,7) = 34.27, p = .001, d = 2.97. This was because it took much longer to select E.2 in Daniel's than in Anna's walls, t(7) = 7.95, P < 0.001, d = 3.97. It also took longer to select E.2 in Daniel's wall relative to D.1, t(7) = 3.13, P = 0.017, d = 1.5. The difference between walls was unexpected. It may relate to the overall change in configuration following the dissimilarity of the layer underneath, as Anna's wall had more bricks in this layer potentially making the placement of the bricks easier. It may also reflect the fact that the Black bricks were the farthest to the left (see sFigure 1) and hence IAI was confound by the extra distance needed to be covered. Though, more importantly the results did not show a chunk boundary effect.