Event Segmentation in Chess

Patric Pförtner (f85787@students.nbu.bg)
Department of Cognitive Science and Psychology, New Bulgarian University,
Sofia, Bulgaria

Penka Hristova (phristova@cogs.nbu.bg)
Department of Cognitive Science and Psychology, New Bulgarian University,
Sofia, Bulgaria

Abstract

How do chess players perceive events in a chess game, as these events unfold in real time? The study builds upon the hierarchical bias hypothesis, stating that observers instinctively segment activities in alignment with a partonomic hierarchy. The alignment effect observed in previous research is replicated, while chess experts outperformed novices. Participants watched chess game videos and identified event boundaries. Data was analysed using discrete, continuous methods, as well as an agreement index. The results aim to deepen our understanding of the cognitive processes involved in chess expertise and event segmentation. They highlight the hierarchical organisation of mental representations in strategic contexts.

Keywords: event segmentation; agreement index; mental representations; event cognition

Introduction

The process of perceiving and understanding our surroundings often involves breaking down experiences into distinct, meaningful units or events, a cognitive process known as event segmentation (Zacks, Braver, Sheridan, Donaldson, Snyder, Ollinger, Buckner & Raichle 2001; Zacks, Speer, Swallow, Braver & Reynolds 2007; Zacks, Tversky & Iyer, 2001). This segmentation allows us to make sense of the continuous flow of information we encounter in our daily lives, enabling us to predict future occurrences and adapt our behaviour accordingly (Zacks, 2020).

Traditionally, researchers have studied event segmentation through tasks where participants watch films, view slideshows, or engage with stories either by reading or listening. However, these tasks aren't just limited to visual or auditory media. In fact, any activity that progresses over time and can be broken down into meaningful segments can be used to study event segmentation (Zacks, 2020). During these tasks, participants mark the boundaries between events by pressing a button (Newtson, 1973). Or in other words, they mark where one event ends and another begins. Before they start, they're usually told to identify event boundaries as either small, detailed segments (referred to as "fine-grained" events) or larger, broader segments (known as "coarse-grained" events).

Perception of event structure for recurring events is influenced by hierarchically organized schemata, known as event schemata (Zacks, Tversky, and Iyer, 2001). Event schemas guide our understanding of stories, memory for events, planning for future activities, and comprehension of our past actions. Moreover, the boundaries of coarse units (broad, general events) tend to align more closely with the boundaries of fine units (specific, detailed events) than would be expected by chance. This alignment effect was found to be mediated by the familiarity of the activity, indicating that the more familiar an individual is with an activity, the more likely they are to perceive its structure in a way that aligns with their existing event schemas and allows them to form expectations about what will happen next (Kurby & Zacks, 2012; Radvansky & Zacks, 2017; Zacks, Speer, Swallow, Braver, & Reynolds, 2007).

When a notable change in important features of the situation, such as new actions, spatial location, objects, causes, and goals, is detected the prediction error tends to increase dramatically. For instance, Loucks and Pechey (2016) discovered that adults are more likely to identify changes in movie clips when these change goal-relevant movement features, compared to when the changes are in features not linked to the action's goals. Furthermore, the ability to process goals, enhances observers' predictive abilities (Zacks, 2020). Analogously, goal-oriented chunking is crucial in areas requiring high levels of expertise, such as chess (Chase & Simon, 1973; De Groot, 1965; Gobet & Simon, 1996). Chess experts, for instance, don't just see individual pieces on the board. Instead, they perceive clusters of pieces, as meaningful units that relate to their strategic goals (Gobet & Simon, 1996). This ability to chunk information based on goals allows them to process the complex configurations of a chess game more efficiently, make superior moves and anticipate their opponents' moves.

Event Segmentation Theory can be considered as a useful framework for measuring this goal-oriented chunking, too. Just as adults are more adept at detecting changes in movie clips when these changes are related to the action's goals (Loucks & Pechey, 2016), chess experts might be more sensitive to changes in the game that are relevant to strategic goals. Hence, experts compared to novices are more likely to detect a significant change in key aspects of the game, such as new unexpected moves, or shifts in strategy. Furthermore, the player's prediction error - the difference between what they expected and what actually happened - should spike. This will force the player to update their current event model, marking an event boundary, just as described in previous research on event boundaries (Loucks, Mutschler & Meltzoff, 2016; Tauzin, 2015). Therefore, chess experts will identify more meaningful events in an event segmentation task for a game of chess.

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

Also, experts' alignment between the fine and coarse segmentation should be superior because they can easily break down goals into subgoals. This process continues recursively until all subgoals are achievable by basic actions or a behavioural primitive (Newell & Simon, 1972). For example, the high-level goal of checkmate may be decomposed into subgoals like controlling the center, developing pieces, protecting the king, planning and executing an attack. Each goal-subgoal relationship manifests as a part-subpart relationship, thereby giving rise to what is termed as a partonomic hierarchy and results in higher agreement between the segmentation tasks (fine vs coarse; Radvansky and Zacks, 2014).

To investigate this, we employed the segmentation procedure outlined by Zacks, Tversky & Iyer (2001) and applied it to the perception of two chess games, each consisting of 80 moves.

Event Segmentation in Chess

Participants were tasked to segment chess games into either the largest and most meaningful units (coarse condition) or smallest and most meaningful units (fine condition). This study investigated the hierarchical bias hypothesis (Zacks, Tversky & Iyer, 2001), suggesting that observers instinctively segment activities in alignment with a partonomic hierarchy. We hypothesised that the fine segmentation of chess experts, due to their extensive knowledge of event schemata and comprehension of goals and sub-goals in chess strategies, would align more closely with coarse segmentation than would be expected by chance. Furthermore, we anticipated that these experts would surpass novices in this aspect. Finally, we utilised the Agreement Index to evaluate its suitability for the segmentation task in the context of chess.

Method

Participants

The initial pool of participants included 30 individuals who played regularly on Chess.com. Participants were recruited through E-Mail or directly on the platform. However, 11 individuals were excluded from the analysis. Among these, five discontinued their participation before reaching the distraction task, while the remaining six ceased participation during the training phase. The final study sample consisted of 19 participants (ages 19–77, M = 44.00 years, SD = 18.32; three females, 16 males) who successfully complied with the study instructions. The expert group consisted of nine individuals (ages 31–77, M = 49.44 years, SD = 17.59; two females, seven males). They started playing chess between the ages of six and 14 (M = 9.89, SD = 2.80). The novice group included 10 individuals (ages 19–71, M = 39.10 years, SD = 18.430; one female, nine males). They started playing chess between the ages of six and 31 (M = 15.30, SD = 8.43). Participants were categorised by two chess rating systems.

Blitz rating: This rating system by Chess.com is applicable to players participating in games with durations exceeding three minutes yet falling short of 10 minutes. In the event of a victory, a player earns points, whereas a defeat leads to a deduction of points from their rating. As the disparity in rating between two players increases, so does the magnitude of points gained or lost. Conversely, a smaller rating difference corresponds to a lesser gain or loss of points. In this study, novices had Blitz ratings ranging from 612 to 1200 (M = 794.90, SD = 176.03), while experts ranged from 1800 to 2169 (M = 1986.11, SD = 96.26).

FIDE rating: An internationally recognised chess rating system by the Federation Internationale des Echecs (FIDE), also known as the World Chess Federation, uses the Elo rating system to rank chess players. Six experts had FIDE ratings between 2001 and 2101 (M = 2032.50, SD = 42.67).

The study began with participants giving an informed consent and completing the preliminary survey. Participants with a Blitz ranking score above 1800 were classified as experts (Gagarin, 2013).

Materials

Videotapes of Chess Games All games were presented on the same chessboard. Each slide was presented for three seconds in both training and the main study, resulting in a total game time of 78 seconds for training and four minutes for a game in the main study. During training, participants viewed a 26-move chess game twice. In contrast, the main study involved two different games, each presenting 80 moves.

In the first training session, the following strategic manoeuvres could be observed: The game initiated with a Queen's Pawn Opening, which transitioned into a Queen's Gambit. The gambit was accepted, leading to a series of exchanges. The game further progressed with a Castling move. Unfortunately, a blunder resulted in a Queen sacrifice, subsequentially leading to a checkmate.

The first game of the main study showcased a variety of strategic motives. It began with a King's Pawn Opening, followed by a Sicilian Defense. The game then transitioned into a Nyezhmetdinov-Rossolimo Attack. Subsequent moves included Castling and an en passant capture, leading to further exchanges. A blunder led to a Queen sacrifice, and the game concluded with a pawn promotion and the immediate check mate event resulting of that promotion.

The second game featured a different set of strategies. The game opened with a King's Pawn Opening, specifically the King's Knight Variation. This was followed by a Ruy López Opening with a Berlin Defense. Both sides performed Castling, and the Rio Gambit was accepted. Despite a series of exchanges and a Queen sacrifice due to a blunder, the game continued without leading to a checkmate.

The selection of specific chess moves, such as the Queen's Gambit and the Nyezhmetdinov-Rossolimo Attack, was deliberate, as these moves are commonly recognized among chess experts. This deliberate choice aimed to facilitate the effective engagement of expert participants in the segmentation task. The games utilised in this study were reconstructed based on a commentary YouTube video featuring Magnus Carlsen. The authors have replicated each move from the video using an open-source chessboard creator. This approach ensured fair use, as the original video was not shown to participants.

Procedure

The experiment was conducted online and participants were provided with a link to access it. The entire script for the experiment was written in Svelte, a modern, component-based JavaScript framework. The experiment was designed to be accessible across various platforms. It was compatible with desktop computers, tablets, and smartphones. The data collected during the study was anonymised.

At the beginning the following information was collected from each participant: (1) their current Chess.com Blitz Ranking, (2) their age, (3) their gender, (4) their proficiency in English, (5) the type of device used to access the experiment (desktop, tablet, or smartphone), (6) their nationality, and (7) the age at which they started playing chess. If applicable, participants could provide FIDE rating.

Then participants were given both general training instructions and specific instructions for each task. The segmentation task involved segmenting a chess game into either the largest and most meaningful units or the smallest and most meaningful units (akin Zacks et al. 2001a; Zacks et al. 2001b). Participants indicated the end of one event and the start of another by pressing a red button beneath the virtual chessboard. A change in border color from black to red confirmed the registration of their response.

Once they confirmed their understanding of the instructions, a three-second countdown began, after which the training task with a chess game of 26 moves was displayed. Upon completion of the game, participants were asked to recall and name as many meaningful units as they could, separating them by commas. After submitting their responses, they proceeded to the next training video. The training was completed with the submission of the training survey, which asked participants about instruction clarity, response accuracy confidence, preferred segmentation task, and video speed.

In the main part of the study, participants watched videos of two chess games. They watched each game twice - once for the 'fine' condition and once for the 'coarse' condition. In between these conditions, they did a distraction task.

Just like in the training, participants were asked to break down the games into either the smallest possible units or the largest units that still were meaningful to them. The instructions before each task told them which type of segmentation task to do (fine or coarse) and which color of chess pieces to focus on. After they confirmed that they understood the task, a three-second countdown appeared. Then, they watched an 80-move chess game. When the game ended, participants were asked to name as many meaningful parts of the game as they could, separating each part with a comma. After participants submitted their responses for the first video, they moved on to the second video. The procedure for the second video was identical to the first one. The tasks (i.e., fine vs. coarse segmentation) were counterbalanced across participants. The order of the videotapes, too.

Between the second and third games, when the type of task changed, participants did a distraction task. In this task, they were asked to come up with a random sequence of numbers between 0 and 9. They had 100 seconds to do this task, and their goal was to come up with at least 100 random numbers. When the time was up, participants were automatically taken to the instruction for the third and fourth games. After they finished the segmentation task for these games, they were asked to fill out a short exit survey. It mirrored the training survey questions about instruction clarity, response accuracy confidence, and video speed and in addition, participants were asked which segmentation was easier instead of preferred.

The experiment typically took a participant 40 minutes.

Results

Participants were tasked with segmenting chess games as they watched them, doing so under both fine and coarse coding instructions. The primary objective of this study was to empirically test the hierarchical bias hypothesis (Zacks, Tversky & Iyer, 2001), which suggests that observers naturally segment activities in a way that aligns with a partonomic hierarchy. The secondary objective was to evaluate whether the Agreement Index is a suitable method of analysis for the segmentation task in the context of chess. For this paper, we will solely report the results of the event segmentation task, excluding surveys and distraction task outcomes.

We will initially present the findings derived from the Agreement Index. Subsequently, we will proceed with the examination of the hierarchical bias hypothesis, employing both discrete and continuous analysis.

Data preparation

Button clicks that appeared on the same slide were deemed as repeats or misplacements and were therefore discarded. This occurred once each in the expert and novice groups. There was one outlier identified in the novice group, as assessed by boxplot in the discrete analysis. It was removed from all further analyses due to the participant's lack of confidence in his replies and understanding of the task, as indicated by survey data. In the discrete analysis, both the observed measures and predicted by chance measures had to undergo a log transformation due to strongly positively skewed data. Furthermore, one participant's data was discarded from the continuous analysis due to its incompatibility with the equation proposed by Zacks, Tversky, and Iyer (2001). It is important to note that this equation necessitates a minimum of two fine segmentations to yield results. However, this participant only provided a single fine segmentation for each chess board examined.

Agreement Index

The agreement index, a measure of how well an individual agrees with a separate group of observers (Sasmita & Swallow, 2022), varied based on the specific chess game video being segmented. As the results will demonstrate, it successfully differentiated actual segmentation from random actions across both coarse and fine data sets.

In the segmentation task, both experts and novices exhibited significant agreement indices, albeit with distinct patterns. Experts demonstrated superior performance in coarse and in fine segmentation. Specifically, the mean Agreement Index for experts was 0.607, with the coarse mean Agreement Index at 0.645 and the fine mean Agreement Index at 0.569. In contrast, novices achieved a mean Agreement Index of 0.506, with a coarse mean Agreement Index of 0.524 and a fine mean Agreement Index of 0.489.

	Experts				
	M	95% C.I.	*t*(8)	*p*	Cohen's *d*
Mean	0.607	[0.807, 3.115]	5.938	< .001	0.573
Fine	0.569	[0.594, 2.640]	4.909	.001	0.510
Coarse	0.645	[0.788, 3.072]	5.84	<.001	0.567
Fine Board 1	0.648	[0.628, 2.713]	5.068	< .001	0.519
Coarse Board 1	0.635	[0.689, 2.849]	5.364	< .001	0.537
Fine Board 2	0.490	[0.474, 2.385]	4.349	.002	0.477
Coarse Board 2	0.654	[0.764, 3.017]	5.727	< .001	0.560
	Novices				
	M	95% C.I.	*t*(8)	*p*	Cohen's *d*
Mean	0.506	[0.448, 2.331]	4.229	.003	0.471
Fine	0.489	[0.394, 2.223]	3.987	.004	0.457
Coarse	0.524	[0.453, 2.341]	4.251	.003	0.472
Fine Board 1	0.449	[0.196, 1.842]	3.117	.014	0.414
Coarse Board 1	0.508	[0.456, 2.347]	4.265	.003	0.473
Fine Board 2	0.529	[0.581, 2.612]	4.847	.001	0.506
Coarse Board 2	0.540	[0.359, 2.153]	3.829	.005	0.449

Table 1: Mean Agreement Indices of Experts and Novices

Furthermore, experts' segmentation indices for both Boards 1 and 2 were significantly different from zero, indicating robust segmentation performance. Specifically, the average coarse Agreement Index was 0.635 for Board 1 and 0.654 for Board 2. Experts exhibited a fine mean Agreement Index of 0.648 for Board 1 and 0.490 for Board 2. Notably, novices achieved a higher fine mean Agreement Index of 0.529 compared to experts on Board 2. The observed discrepancy in the fine Agreement Index for experts on Board 2 is noteworthy.

Alignment effect

The alignment effect from Zacks, Tversky, and Iyer's (2001) study was tested with a discrete analysis, which focuses on Overlaps (Observed Overlaps vs. Overlaps by Chance) representing countable and distinct values, and a continuous analysis, which examines the Average Distance (Actual vs. Predicted by Chance) reflecting a range of possible values. It is hypothesised that experts will exhibit a greater number of Observed Overlaps of fine and coarse segmentation compared to novices. It is also hypothesised that the Average Distance between fine and coarse segmentations will be smaller for experts compared to novices.

In the discrete analyses, we conducted a 2 (Level of Expertise: Experts vs. Novices) × 2 (Overlaps: Observed vs. Predicted by Chance) mixed-design repeated measures ANOVA to assess whether expertise significantly influenced the alignment between the coarse and fine segmentation.

A statistically significant interaction between Overlaps and Level of Expertise was found ($F(1, 16) = 8.176$, $p = .011$, $\omega^2 = .069$). Additionally, there was a significant main effect for Level of Expertise, $F(1, 16) = 7.432$, $p = .015$, $\omega^2 = 0.159$. Post hoc analysis, incorporating the Bonferroni correction, revealed a significant difference in segmentation task performance based on the Level of Expertise, ($M = 0.254$, $SE = 0.093$, $p = .015$). In addition, experts outperformed novices significantly in observed overlaps ($M = 0.377$, $SE = 0.103$, p = .008), too.

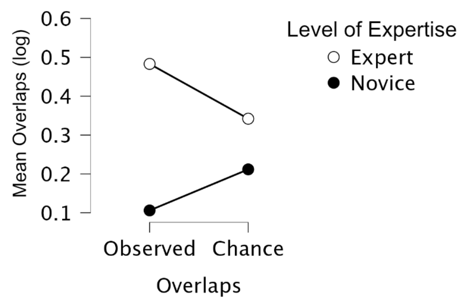

Figure 1: Interaction between Expertise Level and Overlaps in a Segmentation Task.

In the continuous analysis, we performed another mixed-design repeated measures ANOVA with two factors - Level of Expertise (Experts vs. Novices) and Average Distance (Actual vs. Predicted by Chance) - to assess the impact of expertise on alignment effect.

A statistically significant interaction between Average Distance and Level of Expertise was found ($F(1, 15) = 9.974$, $p = .006$, $\omega^2 = 0.118$). This interaction supports our hypothesis, demonstrating that familiarity, as reflected in the expertise of chess players, influences segmentation behaviour. Then, there was a significant main effect for Level of Expertise, $F(1, 15) = 21.271$, $p < .001$, $\omega^2 = 0.388$ and a significant main effect for Average Distance $F(1, 15) = 50.399$, $p < .001$, $\omega^2 = 0.425$.

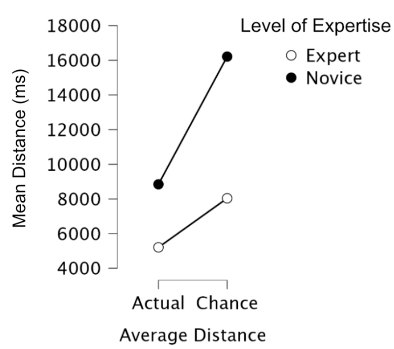

Figure 2: Interaction between Expertise Level and Average Distance in a Segmentation Task.

The post hoc analysis, which included the Bonferroni correction, showed a significant difference in performance based on the Level of Expertise (M = -5907.15, SE = 1280.79, $p < .001$). The performance of experts yielded a smaller Actual Average Distance in comparison to the Chance Average Distance observed in the performance of novices (M = -11007.19, SE = 1468.51, $p < .001$). In a similar manner, the Actual Average Distance was smaller than the Chance Average Distance in the performance of novices (M = -7368.88, SE = 1045.42, $p < .001$). Lastly, a significant difference was found when comparing the Actual Average Distance to the Chance Average Distance (M = -5100.05, SE = 718.40, $p < .001$).

Consistently with the previous findings for alignment effect the boundaries of the coarse units often aligned more closely with the fine units of a chess game than what would be expected by chance. The superior performance of experts in the segmentation tasks, as shown by both discrete and continuous analyses, underscores the role of event schemata and hierarchical segmentation in chess.

Discussion

In our study, we found that chess experts, due to their deep understanding of goals and sub-goals in chess, aligned their fine segmentation more closely with coarse segmentation than novices. Agreement Index effectively captured the group segmentation patterns and differentiated actual segmentation from random actions across both coarse and fine data sets, proving its reliability for future chess expertise research involving segmentation tasks.

While most chess-related studies have concentrated on cognitive processes like memory recall, recognition, and perception of significant game positions, they often overlook the importance of conducting experiments using ongoing games (Chase & Simon, 1973, De Winter, Koelmans, Kokshoorn, Van Der Valk, Vos, Dodou, Eisma, 2023). The ability to analyse and manipulate an ongoing game is crucial, as observing games in progress is a common practice among chess players preparing for tournaments. Our study shows that the segmentation task serves as a dependable method for evaluating segmentation behaviour in settings with limited experimental control.

We demonstrated that although there's no "right" way to perform a segmentation task, the performance isn't arbitrary and a segmentation task can indeed be applied to the domain of chess. This was underlined by successfully replicating Zacks, Tversky, and Iyer's (2001) alignment effect. Our findings revealed that coarse breakpoints were typically closer to the nearest fine breakpoint than expected by chance, as shown by both discrete and continuous analyses. Notably, the results of experts' segmentation indicated the influence of familiarity, event schemas, and successful sub-goal creation on segmentation tasks.

The Event Segmentation Theory can be considered to understand how high-level representations are used and updated during a game, especially when significant changes occur that require the player to revise their current event model. Nonetheless, CHREST, or Chunk Hierarchy and Retrieval Structures, serves already as one cognitive framework in the domain of chess, mimicking human processes of perception, learning, memory, and problem-solving. While CHREST can theoretically incorporate high-level representations, such as the opening from which a position originated and potential plans and moves in the position (Gobet & Simon, 1996), it currently incorporates minimal information of this nature (Lane & Chang, 2017). Our discrete analysis found a statistically significant interaction between Overlaps and Level of Expertise, suggesting that the expertise influences how events are segmented in coarse and fine units. This result is in line with previous findings that highlight the impact of goals and schemas on event segmentation (Loucks & Pechey, 2016; Zacks, 2020). Future investigations into event segmentation in chess could concentrate on elucidating the emergence of these high-level conceptual processes and determining the stages at which they become apparent.

Event segmentation research has traditionally been applied to studies featuring videos of daily activities (Boggia & Ristic, 2015) and visual or written narratives (Bailey, Kurby, Sargent & Zacks, 2017; Sasmita & Swallow, 2022). However, not in strategic games like chess, which requires years of expertise development. Consequently, this study pioneers the exploration of chess as a new domain of expertise in event segmentation research, while also providing a first comparison between chess experts and novices. In doing so, we address a gap in the scientific literature of event segmentation, by incorporating an expertise-related paradigm, a need previously highlighted by Bläsing (2015).

The importance of this method lies in its potential to provide a new, implicit measure of expertise. Traditionally, studies of expertise, particularly in chess, have focused on problem-solving tasks within the domain of expertise (Gobet & Charness, 2018). These tasks often involved right or wrong decisions, like in a problem-solving task or finding the best candidate move task. However, our study, which involved event segmentation, does not necessitate predetermined correct or wrong decisions.Instead, it provides a new look at expert-novice differences, which may push further our understanding of expertise.

The inclusion of expert moves such as the Queen's pawn opening, which may be immediately followed by a Queen's Gambit, enables experts to categorise these moves into single events. This raises the question of whether the segmentation results predominantly reflect the inherent expertise of the players or are influenced by the predefined expert moves embedded within the games. One way to address this limitation is to carefully consider the design of future studies. One viable approach could be to incorporate a control condition where participants are exposed to chess games without expert moves, thereby exposing them to novice-level gameplay. By comparing segmentation performance between conditions (with and without predefined expert moves), future research may assess the extent to which expertise versus the presence of expert moves drives segmentation behaviour. Furthermore, if chess moves were randomly selected or made by novices, any meta-game goals that experts rely on could be absent, potentially altering their event boundaries to resemble those of novices. However, it's important to note that the lack of interaction effects between fine/coarse segmentation and familiarity, as shown in previous studies by Zacks, Tversky, and Iyer's (2001), suggests that expertise in event segmentation may not necessarily be influenced by familiarity with the task or material. McGatlin, Newberry, and Bailey (2018) demonstrated that familiarising younger participants with scripts improved segmentation agreement, whereas their knowledge intervention did not lead to improved segmentation agreement among older adults. Nevertheless, research exploring the effects of conceptual factors on segmentation yields mixed findings, as highlighted by Newberry, Feller, and Bailey (2021).

Finally, future research may investigate two conditions for training novices in segmentation tasks: (1) Exposing novices to problems directly related to segmentation, similar to de Groot's (1946/1965) concept of hierarchical problem-solving in chess. Herein, he proposed that chess thinking is a hierarchical structure where players manage problems by breaking them down into smaller, more manageable subproblems. (2) Allowing novices to observe how experts segment events, followed by describing and labelling the segments themselves. This approach can provide novices with a deeper understanding of the segmentation process beyond simply decomposing a task, potentially leading them to discover the underlying relations between elements as highlighted by Penn, Holyoak, and Povinelli (2008).

While both De Groot (1946/1965) and Gobet and Simon's (1996) template theory emphasize pattern recognition and decomposition, they may underestimate the human capacity for abstract reasoning, generalizing across situations, and flexibly adapting to novel challenges – skills crucial for identifying relations that go beyond surface features. While chess indeed revolves around identifying strong moves, we believe that the segmentation task offers a unique perspective. It allows us to delve into how players perceive and chunk information during a game, providing insights into their cognitive processes beyond mere move selection. This method offers a novel lens to implicitly assess the depth of a player's understanding and expertise in chess, which we argue complements the traditional approaches.

This study employed a post-game recall task for segmentation, which can be susceptible to memory biases. Participants might forget or misremember their in-game segmentation (memory limitations). Additionally, recalled units could be influenced by pre-existing chess knowledge (schema activation) rather than reflecting real-time segmentation. Novice participants, in particular, might fabricate units. If the focus is on memory, offline naming with specific recall prompts might be relevant. If the goal is to understand real-time thought processes, online naming (with careful consideration of cognitive load) might be preferred. Future research may also consider eye-tracking (De Winter et al., 2023) or verbal protocols to capture the segmentation processes.

In conclusion, our study underscores the effectiveness of the segmentation task in assessing meaningful behaviour and contributes to the growing body of research. Further, we highlighted the importance of familiarity and the use of ongoing games in chess-related studies. Our findings highlight the segmentation task's robustness in less controlled environments like the one of an online setting.

References

Bailey, H. R., Kurby, C. A., Sargent, J., & Zacks, J. M. (2017). Attentional focus affects how events are segmented and updated in narrative reading. *Memory & Cognition, 45*(6), 940–955. https://doi.org/10.3758/s13421-017-0707-2

Bläsing, B. E. (2015). Segmentation of dance movement: Effects of expertise, visual familiarity, motor experience and music. *Frontiers in Psychology,* 5. https://doi.org/10.3389/fpsyg.2014.01500

Boggia, J., & Ristic, J. (2015). Social event segmentation. *Quarterly Journal of Experimental Psychology, 68*(4), 731–744. https://doi.org/10.1080/17470218.2014.964738

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*, 55–81. doi: 10.1016/0010-0285(73)90004-2

De Groot, A. D. (1965). *Thought and choice in chess.* The Hague, Netherlands: Mouton.

De Winter, J., Koelmans, T., Kokshoorn, M., Van Der Valk, K., Vos B., Dodou, D., Eisma, Y. (2023). A role of peripheral vision in chess? Evidence from a gaze-contingent method. *Journal of Expertise*, 6(1), 23-38.

Gagarin, G. (2013). FIDE vs Chess.com ratings explained - Chess Forums. Chess.com. Retrieved January 21, 2024, from https://www.chess.com/forum/view/general/fide-ratings-vs-chesscom-ratings-explored

Gobet, F., & Charness, N. (2018). Expertise in chess. In K. A. Ericsson, R. R. Hoffman, A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge handbook of expertise and expert performance* (2nd ed., pp. 597–615). Cambridge University Press. https://doi.org/10.1017/9781316480748.031

Gobet, F., & Simon, H. A. (1996). Templates in Chess Memory: A Mechanism for Recalling Several Boards. *Cognitive Psychology, 31*(1), 1–40. https://doi.org/10.1006/cogp.1996.0011

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, 24, 651–682.

Kurby, C. A., & Zacks, J. M. (2012). Starting from scratch and building brick by brick in comprehension. *Memory & Cognition, 40*(5), 812–826. https://doi.org/10.3758/s13421-011 -0179-8

Lane, D. M., & Chang, Y. (2017). Chess knowledge predicts chess memory even after controlling for chess experience: Evidence for the role of high-level processes. *Memory & Cognition*, 46(3), 337–348. https://doi.org/10.3758/s13421-017-0768-2

Loucks, J., Mutschler, C., & Meltzoff, A. N. (2016). Children's Representation and Imitation of Events: How goal organization Influences 3-Year-Old children's memory for action sequences. *Cognitive Science, 41*(7), 1904–1933. https://doi.org/10.1111/cogs.12446

Loucks, J., & Pechey, M. (2016). Human Action Perception is Consistent, Flexible, and Orientation Dependent.

Perception, 45(11), 1222–1239. https://doi.org/10.1177/0301006616652054

McGatlin KC, Newberry KM, & Bailey HR, (2018). Temporal chunking makes life's events more memorable. *Open Psychology*, 1(1), 94–105.

Newberry, K. M., Feller, D. P., & Bailey, H. R. (2021). Influences of domain knowledge on segmentation and memory. *Memory & cognition*, 49(4), 660–674. https://doi.org/10.3758/s13421-020-01118-1

Newell, A., & Simon, H. A. (1972). *Human problem solving.* Prentice-Hall.

Newtson D. 1973. Attribution and the unit of perception of ongoing behavior. *J. Personal. Soc. Psychol.* 28:28–38

Newtson, D. (1976). Foundations of attribution: The perception of ongoing behavior. *New directions in attribution research, 1*, 223-247.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *The Behavioral and brain sciences*, 31(2), 109–178. https://doi.org/10.1017/S0140525X08003543

Radvansky, G. A., & Zacks, J. M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences, 17*, 133–140. https://doi.org/10.1016/j.cobeha.2017.08.006

Sasmita, K., & Swallow, K. M. (2022). Measuring event segmentation: An investigation into the stability of event boundary agreement across groups. *Behavior Research Methods, 55*(1), 428–447. https://doi.org/10.3758/s13428-022-01832-5

Tauzin, T. (2015). Simple visual cues of event boundaries. *Acta Psychologica, 158*, 8–18. https://doi.org/10.1016/j.actpsy.2015.03.007

Zacks, J. M. (2020). Event perception and memory. *Annual Review of Psychology, 71*(1), 165–191. https://doi.org/10.1146/annurev-psych-010419-051101

Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner, R. L., & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience, 4*(6), 651–655. https://doi.org/10.1038/88486

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: A mind-brain perspective. *Psychological Bulletin, 133*(2), 273–293. https://doi.org/10.1037/0033-2909.133.2.273

Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science, 16*(2), 80–84. https://doi.org/10.1111/j.1467-8721.2007.00480.x

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology: General, 130*(1), 29–58. https://doi.org/10.1037/0096-3445.130.1.29